

DEPTH OF VERTICES WITH HIGH DEGREE IN RANDOM RECURSIVE TREES

LAURA ESLAVA

ABSTRACT. Let T_n be a random recursive tree with n nodes. List vertices of T_n in decreasing order of degree as v^1, \dots, v^n , and write d^i and h^i for the degree of v^i and the distance of v^i from the root, respectively. We prove that, as $n \rightarrow \infty$ along suitable subsequences,

$$\left(d^i - \lfloor \log_2 n \rfloor, \frac{h^i - \mu \ln n}{\sqrt{\sigma^2 \ln n}}\right) \rightarrow ((P_i, i \geq 1), (N_i, i \geq 1)),$$

where $\mu = 1 - (\log_2 e)/2$, $\sigma^2 = 1 - (\log_2 e)/4$, $(P_i, i \geq 1)$ is a Poisson point process on \mathbb{Z} and $(N_i, i \geq 1)$ is a vector of independent standard Gaussians. We additionally establish joint normality for the depths of uniformly random vertices in T_n , which extends results from [8, 14]. The joint holds even if the random vertices are conditioned to have large degree, provided the normalization is adjusted accordingly.

Our results are based on a simple relationship between random recursive trees and Kingman's n -coalescent; a utility that seems to have been largely overlooked.

1. INTRODUCTION

Random recursive trees have been heavily studied since their introduction in 1970 [16], and are closely related to binary search trees, preferential attachment trees and increasing trees in general, see e.g. [4, 10]. In the current work we obtain strong information about the joint law of degrees and depths of maximum and near-maximum degrees and contrast our results to similar results established for linear preferential attachment trees, see [5, 15]. We first recall basic notation and the standard construction of both random recursive trees (RRTs) and linear preferential attachment trees. We use \ln to denote natural logarithms and \log to denote logarithms with base 2.

For $n \geq 1$, let T_n be a random recursive tree with vertex set $[n] = \{1, \dots, n\}$. The standard construction of RRTs, which couples the elements of $(T_n, n \geq 1)$, is the following: Let T_1 be a single vertex labeled 1, which is the root. For $n \in \mathbb{N}$, the tree T_{n+1} is obtained from T_n by adding an edge from a new vertex $n+1$ to a vertex $v_n \in [n]$; the choice of v_n is uniformly random, and is independent for each $n \in \mathbb{N}$. For $v \in [n]$, the depth $h_{T_n}(v)$ is the distance from v to the root in T_n . We write $d_{T_n}(v)$ for the number of children of v in T_n and call this the degree of v in T_n . A particular characteristic of RRTs, as contrasted with other increasing trees e.g. m -ary trees, is that for each $v \in \mathbb{N}$, almost surely $d_{T_n}(v) \rightarrow \infty$ as $n \rightarrow \infty$. Let $\Delta_n = \max_{v \in T_n} d_{T_n}(v)$ be the maximum degree in T_n and let \mathcal{M}_n be the set of vertices in T_n attaining Δ_n .

Linear preferential attachment trees are also constructed recursively, except that the parent v_n of vertex $n+1$ is chosen with probability proportional to the degree of v_n in the current tree. More precisely, for $\alpha > 0$, the linear preferential attachment process $(T_{\alpha,n}, n \geq 1)$ is defined as follows. Let $T_{\alpha,1}$ be a single vertex labeled 1. For $n \in \mathbb{N}$ let $T_{\alpha,n+1}$ be the tree obtained from $T_{\alpha,n}$ by adding an edge from a new vertex $n+1$ to a vertex $v_n \in [n]$. In this case, the $\mathbf{P}(v_n = v)$ is proportional to $\alpha d_{T_{\alpha,n}}(v) + 1$. Note that, in this context, RRTs correspond to the case $\alpha = 0$.

For the linear preferential attachment models, it has been proven that the renormalized maximum degree $n^{-1/(2+1/\alpha)} \Delta_{\alpha,n}$ converges a.s. and in L_p to a positive, finite random variable with absolutely

Date: November 22nd, 2016.

2010 Mathematics Subject Classification. 60C05, 05C80.

Key words and phrases. Kingman's coalescent, random recursive trees, extreme values, near-maximum degrees.

continuous distribution, [15]. Furthermore, the label of the vertex attaining the maximum degree is finite a.s. [5].

For random recursive trees, the picture is quite different. Naturally, if $i < j$ then $d_{T_n}(i)$ stochastically dominates $d_{T_n}(j)$. However, it is unlikely that the root of T_n will attain the maximum degree in T_n . By construction, $d_{T_n}(i)$ is distributed as $\sum_{j=i+1}^n B_j$ where the summands are independent and B_j is distributed as $\text{Bernoulli}(1/j)$. It follows easily that $d_{T_n}(1) = \ln n(1 + o_p(1))$. However, it is known that the maximum degree satisfies $\Delta_n / \log n \rightarrow 1$ a.s. as $n \rightarrow \infty$ [9].

It is also known that the limiting distribution of $\Delta_n - \log n$ is, up to lattice effects, a Gumbel distribution [2, 12]. The latter can be explained since the Gumbel distribution arises as the limiting distribution of the maximum of independent random variables under rather general hypotheses on the laws of such variables. The degrees of vertices in T_n are correlated and are not identically distributed, but between pairs of vertices in T_n the correlation is weak and the Gumbel limit still occurs. This was first shown by Goh and Schmutz [12] using singularity analysis of generating functions. Our approach to RRTs provides a probabilistic explanation of this phenomenon; see [2] for more details.

In [2], Addario-Berry and the author describe the number of high-degree vertices in T_n via the sequence $(d_{T_n}(v) - \lfloor \log n \rfloor, v \in [n])$. They show that, along suitable subsequences, this sequence converges in distribution to a Poisson point process \mathcal{N} in \mathbb{Z} with $\mathbf{E}[|\mathcal{N} \cap [j, \infty)|] = \Theta(2^{-j})$ for all $j \in \mathbb{Z}$.

1.1. Statement of results. This work provides a detailed description of the degrees and depths of high-degree vertices in T_n . In particular we show that the number of vertices attaining the maximum degree is random and their depths are independent and asymptotically normal. Write $\mu = 1 - (\log e)/2$ and $\sigma^2 = 1 - (\log e)/4$.

Theorem 1.1. *For each $\varepsilon \in [0, 1]$, there exists a positive integer-valued random variable M_ε such that, for any increasing sequence of integers $(n_l, l \geq 1)$ for which $\log n_l - \lfloor \log n_l \rfloor \rightarrow \varepsilon$ as $l \rightarrow \infty$, then $|\mathcal{M}_{n_l}|$ converges to M_ε in distribution, and*

$$\left(\frac{h_{T_{n_l}}(v) - \mu \ln n_l}{\sqrt{\sigma^2 \ln n_l}}, v \in \mathcal{M}_{n_l} \right) \xrightarrow{\mathcal{L}} (N_i, 1 \leq i \leq M_\varepsilon),$$

where N_i are independent standard Gaussian variables.

We remark that Theorem 1.1 implies that maximum-degree vertices of RRTs are constantly changing along the process $(T_n, n \geq 1)$.

Our main result gives a more general description of the depths of all vertices in T_n , indexed in decreasing order of their degrees. List vertices of T_n in decreasing order of degree as v^1, \dots, v^n ; here we break ties between vertices with the same degree by ordering them uniformly at random. Write d^i and h^i for the degree and depth of v^i , respectively. Let \mathcal{P} be a Poisson point process in \mathbb{R} with intensity $\lambda(x) = 2^{-x} \ln 2$. Then for $i \geq 1$, let P_i be the i -th largest point of \mathcal{P} so $|\mathcal{P} \cap [P_i, \infty)| = i$ and $|\mathcal{P} \cap (P_i, \infty)| = i - 1$. This ordering is well defined as $|\mathcal{P} \cap [0, \infty)| < \infty$ almost surely.

Theorem 1.2. *Let N_i be independent standard Gaussian variables, $i \in \mathbb{N}$. For each $\varepsilon \in [0, 1]$ and for any increasing sequence of integers $(n_l, l \geq 1)$ for which $\log n_l - \lfloor \log n_l \rfloor \rightarrow \varepsilon$ as $l \rightarrow \infty$, then*

$$\left(d^i - \lfloor \log n_l \rfloor, \frac{h^i - \mu \ln n_l}{\sqrt{\sigma^2 \ln n_l}} \right) \xrightarrow{\mathcal{L}} ((\lfloor P_i + \varepsilon \rfloor, i \geq 1), (N_i, i \geq 1)).$$

The condition on the subsequence n_l in Theorems 1.1 and 1.2 is due to a lattice effect on the law of $(\lfloor P_i + \varepsilon \rfloor, i \geq 1)$ caused by the fact that degrees are integer-valued.

Our last result provides information about vertices with degree near $a\Delta_n$ for fixed $a \in [0, 1]$. For $a \in [0, 1]$, let $\mu_a = 1 - (a \log e)/2$ and $\sigma_a^2 = 1 - (a \log e)/4$; note that $\mu = \mu_1$ and $\sigma = \sigma_1$.

Theorem 1.3. Fix $k \in \mathbb{N}$ and let $(u_i, i \in [k])$ be k distinct vertices in T_n chosen uniformly at random. For every $(a_1, \dots, a_k) \in [0, 1]^k$ and $(b_1, \dots, b_k) \in \mathbb{Z}^k$, the conditional law of

$$\left(\frac{h_{T_n}(u_i) - \mu_{a_i} \ln n}{\sqrt{\sigma_{a_i}^2 \ln n}}, i \in [k] \right),$$

given that $d_{T_n}(u_i) \geq \lfloor a_i \log n \rfloor + b_i$ for all $i \in [k]$, converges to the law of k independent standard Gaussian variables.

Note that the case $b_i = a_i = 0$ for all $i \in [k]$ of Theorem 1.3 involves no conditioning, and thus yields the joint distribution for the depths of k uniformly random vertices in T_n . This extends the results of the papers [8, 14] where the case for $k = 1$, $a_1 = b_1 = 0$ of Theorem 1.3 is established. These results were obtained in the context of analyzing the *insertion depth*, $h_{T_n}(n)$ of RRTs, important for the analysis of data structures in computer science.

Theorem 1.1 is a quite straightforward consequence of Theorem 1.2, whose proof relies essentially on Theorem 1.3. The proof of Theorem 1.3 exploits the relation between degrees and depths of vertices in a different random tree $T^{(n)}$ whose shape has the same law as that of T_n . This alternative tree $T^{(n)}$ is constructed through Kingman's coalescent, as described in Section 2.1. A binary tree representation of Kingman's coalescent had been previously used to study a data structure known as union-find trees, [8]. Pittel mentions the connection between the results of [8] and the height of RRTs in [17]. However, although the connection between Kingman's coalescent and random recursive trees had been observed, prior to our previous work with Addario-Berry [2], its utility in studying vertex degrees seems to have gone unremarked.

1.2. The point process in Theorem 1.2. In this section we briefly explain how we use the method of moments (e.g., see [13] Section 6.1) to obtain the limiting distribution of a sequence of (marked) point processes. In particular, we present an alternative characterization of the processes involved in Theorem 1.2. Although this change of perspective requires the introduction of further notation, the problem of establishing Theorem 1.2 becomes, in fact, more tractable.

We start by considering the unmarked processes $(d^i - \lfloor \log n \rfloor, i \in [n])$ and $\mathcal{P} = (P_i, i \geq 1)$. Define, for each $n \in \mathbb{N}$, $\varepsilon_n = \log n - \lfloor \log n \rfloor$. We consider a fixed $\varepsilon \in [0, 1]$ and increasing sequence n_l such that $\varepsilon_{n_l} \rightarrow \varepsilon$ as $l \rightarrow \infty$.

For $j \in \mathbb{Z}$, we define the following counting measures of the sequence $\mathcal{P}^\varepsilon = (\lfloor P_i + \varepsilon \rfloor, i \geq 1)$;

$$\begin{aligned} X_j &= \#\{i \geq 1 : \lfloor P_i + \varepsilon \rfloor = j\}, \\ X_{\geq j} &= \#\{i \geq 1 : \lfloor P_i + \varepsilon \rfloor \geq j\}. \end{aligned}$$

Note that $X_j \stackrel{\mathcal{L}}{=} \text{Poi}(2^{-j+\varepsilon-1})$ and $X_{\geq j} \stackrel{\mathcal{L}}{=} \text{Poi}(2^{-j+\varepsilon})$; in particular, the number of points of \mathcal{P}^ε on any interval $[j, \infty]$ is finite almost surely. Therefore, \mathcal{P}^ε is characterized by the collection of joint distributions $(X_{j'}, \dots, X_{j-1}, X_{\geq j})$ for any integers $j' < j$; see e.g. Section 3.1 of [6] and Section 9.2 of [7]. Similarly, the collection of the joint distribution of the variables

$$\begin{aligned} X_j^{(n)} &= \#\{v \in [n] : d_{T_n}(v) = \lfloor \log n \rfloor + j\}, \\ X_{\geq j}^{(n)} &= \#\{v \in [n] : d_{T_n}(v) \geq \lfloor \log n \rfloor + j\} \end{aligned}$$

characterizes the law of the sequence $(d^i - \lfloor \log n \rfloor, i \geq 1)$. Finally, to prove that $(d^i - \lfloor \log n_l \rfloor, i \geq 1) \xrightarrow{\mathcal{L}} \mathcal{P}^\varepsilon$ it suffices to show that, for all $j' < j \in \mathbb{Z}$

$$(1) \quad (X_{j'}^{(n_l)}, \dots, X_{j-1}^{(n_l)}, X_{\geq j}^{(n_l)}) \xrightarrow{\mathcal{L}} (X_{j'}, \dots, X_{j-1}, X_{\geq j}).$$

Next, for any $r \in \mathbb{R}$ and $a \in \mathbb{N}$, let $(r)_a = r(r-1) \cdots (r-a+1)$ and set $(r)_0 = 1$. Recall that, if $X \stackrel{\mathcal{L}}{=} \text{Poi}(\lambda)$, $\mathbf{E}[(X)_a] = \lambda^a$ for all integers $a \geq 0$. Now, using the method of moments, the following estimates imply (1).

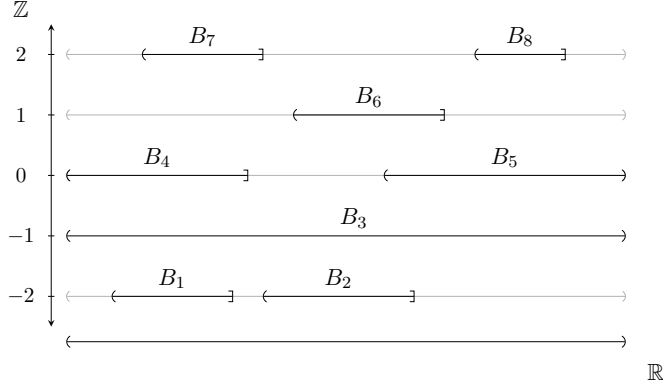


FIGURE 1. An example of a (K', K) -canonical set. In this case, $K' = 6$, $K = 8$ and $j_1 = -2$, $j_8 = 2$.

Proposition 1.4 (Proposition 2.1, [2]). *For all $c \in (0, 2)$ and $A \in \mathbb{N}$ there is $\beta = \beta(c, A) > 0$ such that the following holds. If $j' = j'(n)$ and $j = j(n)$ are integer-valued functions with $0 < j' + \log n < j + \log n < c \ln n$, then uniformly over non-negative integers $a_{j'}, \dots, a_j$ with $a_{j'} + \dots + a_j = A$, we have*

$$\mathbf{E} \left[(X_{\geq j}^{(n)})_{a_j} \prod_{j' \leq k < j} (X_k^{(n)})_{a_k} \right] = (2^{-j+\varepsilon_n})^{a_j} \prod_{j' \leq k < j} \left(2^{-(k+1)+\varepsilon_n} \right)^{a_k} (1 + o(n^{-\beta})).$$

Marked point processes are, in fact, point processes in a larger space; thus, the same approach can be used when we add the information of the depths $((h^i - \mu \ln n) / \sqrt{\sigma^2 \ln n}, i \geq 1)$ and the marks $(N_i, i \geq 1)$. Let us define subsets of $\mathbb{Z} \times \mathcal{B}(\mathbb{R})$ that will help us define the FDDs of our marked point processes; see Figure 1 for an example. It suffices to consider the set

$$\mathcal{B}_I = \{(-\infty, b], (a, b], (a, \infty); -\infty < a < b < \infty\}.$$

Definition 1.5. *Fix positive integers $K' < K$. If the pairs $(j_k, B_k) \in \mathbb{Z} \times \mathcal{B}_I$, $k \in [K]$, satisfy*

- (1) $j_1 \leq j_2 \leq \dots \leq j_{K'} < j = j_{K'+1} = \dots = j_K$ and
- (2) for all $1 \leq k < l \leq K$, if $j_k = j_l$ then $B_k \cap B_l = \emptyset$;

then we say that $((j_k, B_k), k \in [K])$ is a (K', K) -canonical FDD sequence.

Also, for $(j, B) \in \mathbb{Z} \times \mathcal{B}_I$, let

$$X_j(B) = \#\{i \geq 1 : \lfloor P_i + \varepsilon \rfloor = j, N_i \in B\}$$

$$X_j^{(n)}(B) = \#\left\{v \in [n] : d_{T_n}(v) = \lfloor \log n \rfloor + j, \frac{h_{T_n}(v) - \mu_1 \ln n}{\sqrt{\sigma_1^2 \ln n}} \in B\right\};$$

and let $X_{\geq j}(B)$, $X_{\geq j}^{(n)}(B)$ be defined accordingly.

Now the convergence in distribution of point processes is equivalent to the convergence of its finite dimensional distributions (FDD); see [7, Theorem 11.1.VII]. In our case, this leads to the following lemma.

Lemma 1.6. *The following are equivalent.*

- a) As $l \rightarrow \infty$,

$$\left(d^i - \lfloor \log n_l \rfloor, \frac{h^i - \mu \ln n_l}{\sqrt{\sigma^2 \ln n_l}} \right) \xrightarrow{\mathcal{L}} ((\lfloor P_i + \varepsilon \rfloor, i \geq 1), (N_i, i \geq 1)),$$

b) For every (K', K) -canonical FDD sequence $((j_k, B_k), 1 \leq k \leq K)$ as $l \rightarrow \infty$,

$$(X_{j_1}^{(n_l)}(B_1), \dots, X_{j_{K'}}^{(n_l)}(B_{K'}), X_{\geq j_{K'+1}}^{(n_l)}(B_{K'+1}), \dots, X_{\geq j_K}^{(n_l)}(B_K)) \\ \xrightarrow{\mathcal{L}} (X_{j_1}(B_1), \dots, X_{j_{K'}}(B_{K'}), X_{\geq j_{K'+1}}(B_{K'+1}), \dots, X_{\geq j_K}(B_K)).$$

Let Φ denote the measure of a standard Gaussian variable; that is $\Phi(A) = \int_A e^{-x^2/2} dx / \sqrt{2\pi}$ for any $A \subset \mathbb{R}$.

Fact 1.7. For all $j \in \mathbb{Z}$ and $B \subset \mathbb{R}$, $X_j(B) \stackrel{\mathcal{L}}{=} \text{Poi}(2^{-j+\varepsilon-1}\Phi(B))$ and $X_{\geq j}(B) \stackrel{\mathcal{L}}{=} \text{Poi}(2^{-j+\varepsilon}\Phi(B))$; additionally, for any (K', K) -canonical FDD sequence, the variables

$$(X_{j_1}(B_1), \dots, X_{j_{K'}}(B_{K'}), X_{\geq j_{K'+1}}(B_{K'+1}), \dots, X_{\geq j_K}(B_K))$$

are independent.

Below is the more general form of moment estimation which is required to obtain Theorem 1.2.

Proposition 1.8. Fix $c \in (0, 2)$ and $M \in \mathbb{N}$. Let $j = j(n)$ and $j' = j'(n)$ are integer-valued functions with $0 \leq j'(n) + \log n < j(n) + \log n < c \ln n$, and let $K \in \mathbb{N}$ and non-negative integers $(a_k, k \in [K])$ $K \in \mathbb{N}$ be such that $\sum_{k \in [K]} a_k = M$. Then uniformly over (K', K) -canonical sequences $((j_k, B_k), 1 \leq k \leq K)$ with $j' = j_1$ and $j = j_K$, we have

$$\mathbf{E} \left[\prod_{k=1}^{K'} \left(X_{j_k}^{(n)}(B_k) \right)_{a_k} \prod_{k=K'+1}^K \left(X_{\geq j}^{(n)}(B_k) \right)_{a_k} \right] \\ = \prod_{k=1}^{K'} \left(2^{-j_k+\varepsilon_n-1} \Phi(B_k) \right)^{a_k} \prod_{k=K'+1}^K \left(2^{-j'+\varepsilon_n} \Phi(B_k) \right)^{a_k} (1 + o(1)).$$

Proof of Theorem 1.2 (assuming Proposition 1.8). Fix $\varepsilon \in [0, 1]$ and let n_l be an increasing sequence with $\varepsilon_{n_l} \rightarrow \varepsilon$ as $l \rightarrow \infty$. Let $K' < K$ and $((j_k, B_k), 1 \leq k \leq K)$ be a fixed (K', K) -canonical FDD sequence. Set $c = 3/2$, which implies for n large enough that $0 \leq j_1 + \log n < j_K + \log n < c \ln n$. Thus, Proposition 1.8 implies, by the method of moments, that the first condition in Lemma 1.6 is satisfied for each (K, K') -canonical FDD sequence. This completes the proof of Theorem 1.2. \square

We briefly explain a key ingredient to proving Proposition 1.8. Note that each $X_j^{(n)}(B)$ and $X_{\geq j}^{(n)}(B)$ is a sum of indicator variables. Therefore, the expectations of their factorial moments are reduced to a sum of probabilities as follows: for each $S \subset [n]$, collection $B_j \subset \mathcal{B}_I$ and sequence $m_j < 2 \ln n$,

$$\mathbf{P}(\deg_{T_n}(v_j) \geq m_j, h_{T_n}(v_j) \in B_j, v_j \in S) \\ (2) \quad = \mathbf{P}(\deg_{T_n}(v_j) \geq m_j, v_j \in S) \mathbf{P}(h_{T_n}(v_j) \in B_j, v_j \in S \mid \deg_{T_n}(v_j) \geq m_j, v_j \in S).$$

The first factor in (2) has been analyzed in [2]; the result we need from that paper is restated below as Proposition 5.4. The second factor in (2) is bounded in Theorem 2.5.

We now turn to describing Kingman's coalescent.

2. A KINGMAN'S COALESCENT APPROACH

The connection between RRTs and Kingman's coalescent is central to understanding the close relation between degree and depth of vertices reflected in Theorem 1.3, which is key to the proofs of Theorems 1.1 and 1.2. Therefore, we briefly sketch the role that Theorem 1.3 plays in the proof of Theorem 1.2. In Section 2.1 we define the tree $T^{(n)}$, after which we discuss the contents of the remainder of the paper.

2.1. Kingman's coalescent process. In this section we give a representation of Kingman's coalescent in terms of labeled forests and connect this with RRTs. For a general description of Kingman's coalescent, see [3, Chapter 2]; the construction below is based on that given in [1]. For the remainder of the paper, all trees are rooted and we use $r(t)$ to denote the root of tree t . We write $V(t)$ and $E(t)$ for the sets of vertices and edges of t , respectively. By convention, we assume that edges of a tree t are directed towards $r(t)$ and an edge directed from u to v is denoted by uv . If t has n vertices, we say that t has size n ; we also write $d_t(v)$ and $h_t(v)$ for the degree and depth of vertex v in t .

A rooted labeled tree t is *increasing* if its labels are increasing along root-to-leaf paths. Let us write $\mathcal{I}_n = \{t : t \text{ is increasing, } V(t) = [n]\}$ to denote the set of increasing trees on $[n]$. It is not difficult to see that T_n is a uniformly random element of \mathcal{I}_n and that $|\mathcal{I}_n| = (n-1)!$.

A forest f is a set of trees whose vertex sets are pairwise disjoint. Denote by $V(f)$ and $E(f)$, respectively, the union of the vertex and edge sets of the trees contained in f . For each $n \geq 1$, we consider the set of forests $\mathcal{F}_n = \{f : V(f) = [n]\}$ with vertex labels $[n]$. An n -chain is a sequence $C = (f_n, \dots, f_1)$ of elements of \mathcal{F}_n if for $1 < i \leq n$, f_{i-1} is obtained from f_i by adding an edge connecting two of the roots in f_i . In particular, f_n contains n one-vertex trees, and f_1 contains exactly one tree denoted by $t_C \in \mathcal{F}_n$.

For an n -chain $(f_n, \dots, f_1) \in \mathcal{CF}_n$ and $1 \leq i \leq n$, we write $f_i = \{t_1^{(i)}, \dots, t_i^{(i)}\}$, ordering of the trees is in increasing order of their smallest-labeled vertex.

Definition 2.1. *The following constructs Kingman's n -coalescent as a random n -chain $\mathbf{C} = (F_n, \dots, F_1)$.*

For each $1 < i \leq n$, choose $\{a_i, b_i\} \subset \{\{a, b\} : 1 \leq a < b \leq i\}$ independently and uniformly at random; also let $(\xi_i, i \in [n-1])$ be a sequence of independent Bernoulli(1/2) random variables.

For $1 \leq i < n$, F_i is obtained from F_{i+1} as follows. Add an edge e_i between the roots of $r(T_{a_{i+1}}^{(i+1)})$ and $r(T_{b_{i+1}}^{(i+1)})$; direct e_i towards $r(T_{a_{i+1}}^{(i+1)})$ if $\xi_i = 1$, and towards $r(T_{b_{i+1}}^{(i+1)})$ otherwise. Then F_i contains the new tree and the remaining $i-1$ unaltered trees from F_{i+1} .

For an example of the process see Figure 2.

Lemma 2.2. *Kingman's n -coalescent \mathbf{C} is uniformly random in \mathcal{CF}_n , the set of n -chains.*

Proof. Any $(f_n, \dots, f_1) \in \mathcal{CF}_n$ is determined by the order in which the edges of t_C are added. For each $2 \leq i < n$, there are $(i+1)i$ possible oriented edges between the roots in f_{i+1} and only one of them is $e \in E(f_i) \setminus E(f_{i+1})$. Thus,

$$\mathbf{P}((F_n, \dots, F_1) = (f_n, \dots, f_1)) = \prod_{k=1}^{n-1} \mathbf{P}(F_k = f_k | F_j = f_j, k < j \leq n) [n!(n-1)!]^{-1}.$$

This expression holds for all $(f_n, \dots, f_1) \in \mathcal{CF}_n$, so the result follows. \square

Let e_{n-1}, \dots, e_1 be the edges of t_C ordered as they were added to the chain C . That is, $e_i \in E(F_i)$ while $e_i \notin E(F_{i+1})$ for all $1 \leq i < n$. Now, write $e_i = v_i w_i$. Let $\sigma_C : V(t_C) \rightarrow [n]$ be defined as $\sigma_C(r(t_C)) = 1$ and for each $e_i = v_i w_i \in E(t_C)$,

$$\sigma_C(v_i) = i + 1.$$

This is well defined as all edges are directed towards the root, so $v_i \neq v_j$ for all $i, j \in [n-1]$. Note that for each $1 \leq i < n$, e_i is directed towards the root of the new tree in f_i . Thus, the labels $\{\sigma_C(v), v \in [n]\}$ decrease along leaf-to-root paths in t_C . As a consequence, we obtain an increasing tree by relabeling the vertices of t_C using σ_C .

Proposition 2.3. *For each $C = (f_n, \dots, f_1) \in \mathcal{CF}_n$, relabel the vertices in t_C with σ_C to obtain $\phi(C) \in \mathcal{I}_n$. Then the law of $\phi(\mathbf{C})$ is that of a RRT of size n .*

Proof. From the argument in the proof of Lemma 2.2, we have that $|\mathcal{CF}_n| = n!(n-1)!$. Next, we show that ϕ is onto and, additionally, an $n!$ -to-1 mapping. Thus ϕ preserves the uniform measure from \mathcal{CF}_n to \mathcal{I}_n .

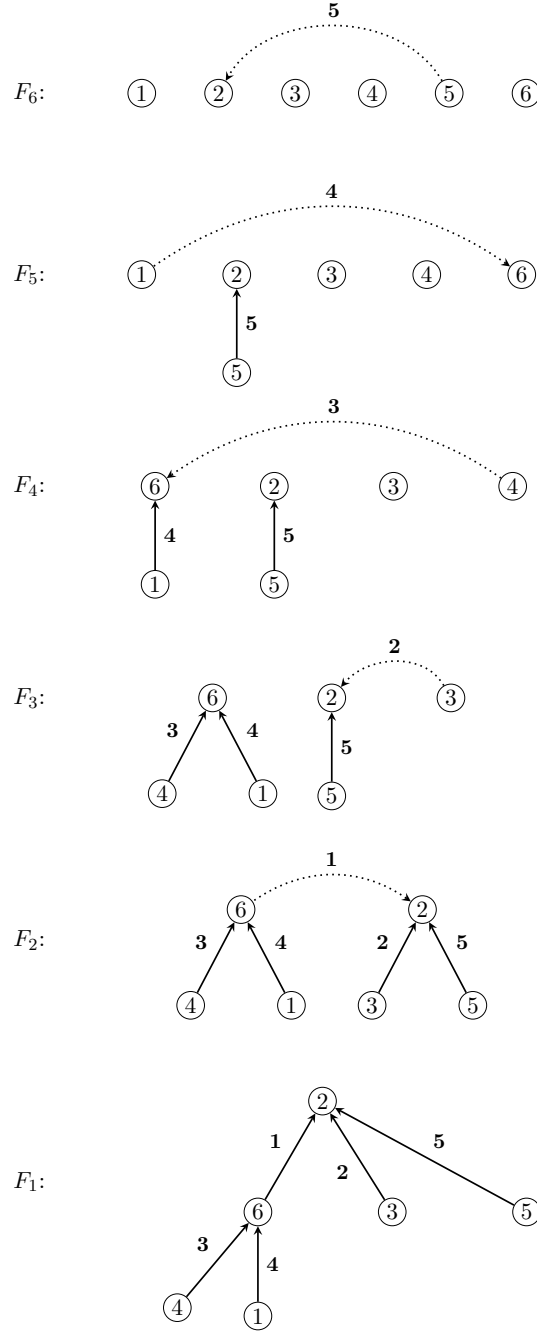


FIGURE 2. An example of Kingman's n -coalescent $\mathbf{C} = (F_n, \dots, F_1)$ for $n = 6$. For $1 < i \leq n$, we present the edge $E(F_{i-1}) \setminus E(F_i)$ with a dotted line in F_i . Edges are marked with the labels ρ_C ; $n - \rho_C(e)$ is the first forest where e is present. In this case, $\xi_6 = \xi_4 = \xi_3 = 1$, $\xi_5 = \xi_2 = 0$ and $\{a_5, b_5\} = \{2, 5\}$, $\{a_4, b_4\} = \{1, 5\}$, $\{a_3, b_3\} = \{1, 4\}$, $\{a_2, b_2\} = \{2, 3\}$, $\{a_1, b_1\} = \{1, 2\}$.

Fix an increasing tree $t \in \mathcal{I}_n$. Every vertex $j > 1$ has outdegree 1 in t , thus we write uniquely define $v_j \in V(t)$ such that $ju_j \in E(t)$. For each $1 < j \leq n$, let $e_{j-1} = ju_j$. Consider an n -chain $C = (f_n, \dots, f_1)$ defined as follows. Let $f_n \in \mathcal{F}_n$ have no edges, and for each $1 \leq i < n$, construct f_i from f_{i+1} by adding the edge e_i . It is easy to see that C satisfies $\sigma_C(i) = i$ for all $i \in [n]$ and t_C . Therefore $\phi(C) = t$, showing that ϕ is onto.

Now, consider $C \in \mathcal{CF}_n$ such that $\phi(C) = t$. For each permutation $\pi : [n] \rightarrow [n]$, let C_π be the n -chain obtained from $C = (f_n, \dots, f_1)$ by applying π to each of the labels of $V(f_i)$, $i \in [n]$. The mapping ϕ does not depend of the vertex labels in C , but on the order in which edges are added; therefore, $\phi(C) = \phi(C_\pi)$ for all permutations π . This shows that $|\phi^{-1}(t)| \geq n!$ for any $t \in \mathcal{I}_n$, completing the proof. \square

For each n , let \mathbf{C} be a Kingman's n -coalescent and let $T^{(n)} = t_{\mathbf{C}}$. Since $\phi(\mathbf{C})$ only relabels vertices in T_C , it follows that the shape of the tree is preserved; and so are the degrees and depths of the vertices. That is, as multisets,

$$\{(deg_{T^{(n)}}(v), h_{T^{(n)}}(v))\}_{v \in [n]} = \{(deg_{\phi(\mathbf{C})}(v), h_{\phi(\mathbf{C})}(v))\}_{v \in [n]}.$$

Moreover, for each $t \in \mathcal{I}_n$ the set $\phi^{-1}(t)$ can be indexed by permutations on $[n]$. This directly implies the following key corollary of Proposition 2.3.

Corollary 2.4. *For all $n \in \mathbb{N}$,*

$$((d_{T^{(n)}}(i), h_{T^{(n)}}(i)), i \in [n]) = ((d_{T_n}(\sigma(i)), h_{T_n}(\sigma(i))), i \in [n]);$$

where σ is a uniformly random permutation of $[n]$ and is independent of T_n . Consequently, the following equality in distribution holds jointly for all $i \in \mathbb{Z}$ and $j \in \mathbb{N}$,

$$|\{v \in [n] : d_{T_n}(v) = i, h_{T_n}(v) = j\}| = |\{v \in [n] : d_{T^{(n)}}(v) = i, h_{T^{(n)}}(v) = j\}|$$

Proof. For any $n \in \mathbb{N}$, let \mathcal{P}_n be the set of permutations on $[n]$. For any n -chain $C = (f_n, \dots, f_1)$ let $\varphi(C) = (\phi(C), \sigma_C)$. Then $\varphi : \mathcal{CF}_n \rightarrow \mathcal{I}_n \times \mathcal{P}_n$ is a bijection and the result follows. \square

2.2. Conditional depths of high-degree vertices. In this section we provide a heuristic for the approach we use to study the conditional distributions involved in Theorem 2.5 below, which is equivalent to Theorem 1.3, and also outline the remainder of the paper.

Fix $n \in \mathbb{N}$ and consider Kingman's n -coalescent $\mathbf{C} = (F_n, \dots, F_1)$. For each vertex $v \in [n]$ and $1 \leq i \leq n$, let $T_i(v)$ be the tree in F_i that contains v . We use $d_{F_i}(v)$ and $h_{F_i}(v)$ to denote the degree and depth of v in $T_i(v)$. Recall that $T^{(n)} = t_{\mathbf{C}}$ is the unique tree in F_1 ; for simplicity, we use $d_n(v)$ and $h_n(v)$ for the degree and depth of vertices in $T^{(n)}$.

Theorem 2.5. *Fix $k \in \mathbb{N}$. For any $(a_1, \dots, a_k) \in [0, 1]^k$ and $(b_1, \dots, b_k) \in \mathbb{Z}^k$, the conditional law of*

$$\left(\frac{h_n(i) - \mu_{a_i} \ln n}{\sqrt{\sigma_{a_i}^2 \ln n}}, i \in [k] \right),$$

given that $d_n(i) \geq \lfloor a_i \log n \rfloor + b_i$, $i \in [k]$, converges to the law of k independent standard Gaussian variables.

Remark 2.6 (Proof of Theorem 1.3). *By Corollary 2.4, Theorem 1.3 follows from Theorem 2.5.*

In this section we give a heuristic for Theorem 2.5, when $k = 1$. First, we analyze the case $m_1 = m_1(a_1, b_1, n) = \lfloor a_1 \log n \rfloor + b_1 \leq 0$, in which $\{d_n(1) \geq m_1\}$ occurs; and second $m_1 > 0$. Finally, we discuss the obstacles in treating several vertices, that is, when $k \geq 2$.

We next define indicator functions $(s_{i,v}, 2 \leq i \leq n)$ and the *selection set* $\mathcal{S}_n(v)$ as follows, let $s_{i,v}$ be the indicator that $T_i(v) \in \{T_{a_i}^{(i)}, T_{b_i}^{(i)}\}$; that is, $s_{i,v} = 1$ when $T_i(v) \in F_i$ is chosen to be merged and form a larger tree in F_{i-1} , and otherwise $s_{i,v} = 0$. Now we set

$$\mathcal{S}_n(v) = \{2 \leq i \leq n : s_{i,v} = 1\}.$$



FIGURE 3. For $1 < i \leq n$ let $r_i = r(T_i(v))$ and suppose $i \in \mathcal{S}_n(v)$. If e_i is directed towards r_i , then the degree of r_i increases by one in F_{i-1} . If e_i is directed outwards r_i , then the depth of each $u \in T_i(v)$ increases by one in F_{i-1} .

The selection set $\mathcal{S}_n(v)$ keeps track of each time i where $T_i(v)$ merges. The choice of trees to be merged at each step is both independent and uniform. Thus, for fixed $v \in [n]$, the variables $(s_{i,v}, 2 \leq i \leq n)$ are independent Bernoulli random variables, with $\mathbf{E}[s_{i,v}] = 2/i$. This implies that $\mathbf{E}[|\mathcal{S}_n(v)|] = \sum_{i=1}^n \frac{2}{i} = 2 \ln n + O(1)$ and $\mathbf{Var}[|\mathcal{S}_n(v)|] = \sum_{i=1}^n \left(\frac{2}{i} - \frac{4}{i^2}\right) = 2 \ln n + O(1)$. It is straightforward to see that the Lindenberg conditions are satisfied by $|\mathcal{S}_n(v)|$ and thus, the following holds for any vertex $v \in \mathbb{N}$,

$$(3) \quad \frac{|\mathcal{S}_n(v)| - 2 \ln n}{\sqrt{2 \ln n}} \xrightarrow{\mathcal{L}} N;$$

as $n \rightarrow \infty$ and where N is a standard Gaussian variable. Moreover, Bernstein's inequalities (see, e.g. [13, Theorem 2.8 and (2.9)]) yield that, for any $\delta > 0$,

$$(4) \quad \mathbf{P}(||\mathcal{S}_n(v)| - \mathbf{E}[|\mathcal{S}_n(v)|]| > \delta \mathbf{E}[|\mathcal{S}_n(v)|]) = o(1).$$

Now, consider the indicator random variables $(\kappa_{i,v}, 2 \leq i \leq n)$ where $\kappa_{i,v} = 1$ precisely when $s_{i,v} = 1$ and the edge added to F_i is directed outwards of $r(T_i(v))$. The latter condition depends only on ξ_i and thus $\mathbf{E}[\kappa_{i,v}] = 1/i$. Recall that e_i is the edge added to F_{i+1} to obtain F_i . If e_i is directed towards $r(T_{i+1}(v))$, the degree of $r(T_{i+1}(v))$ increases by one in F_i . Otherwise, e_i is directed outwards $r(T_{i+1}(v))$ and all vertices in $T_{i+1}(v)$ increase their depth by one in F_i . Therefore $h_{F_j}(v) = \sum_{i=j+1}^n \kappa_{i,v}$, and in particular

$$h_n(v) = \sum_{i=2}^n \kappa_{i,v}.$$

Similarly to (3), it follows that $(h_n(v) - \ln n)/\sqrt{\ln n}$ converges in distribution to a standard Gaussian variable; this already solves the case when $m_1 \leq 0$. However, such arguments cannot be directly applied to the case when $m_1 > 0$ or when $k \geq 2$. We next describe a slightly different proof that $(h_n(v) - \ln n)/\sqrt{\ln n}$ is asymptotically normal, which we later extend to cover the general case of Theorem 2.5.

The direction of the edge e_i is determined by a Bernoulli(1/2), independent of the choice of trees to be merged. Thus, we have the following distributional equality,

$$(5) \quad h_n(1) \stackrel{\mathcal{L}}{=} \text{Bin}(|\mathcal{S}_n(1)|, 1/2).$$

Now, from (3), it follows that there exist random variables $X_n \xrightarrow{\mathcal{L}} N$ such that

$$S_n = |\mathcal{S}_n(1)| = 2 \ln n + X_n \sqrt{2 \ln n}.$$

Similarly, the central limit theorem allows us to write $\text{Bin}(2m, 1/2) = m + \frac{Y_m}{2}\sqrt{2m}$ with $Y_m \xrightarrow{\mathcal{L}} N'$, N' a standard Gaussian variable. We then have

$$\begin{aligned} \text{Bin}(S_n, 1/2) &= \frac{S_n}{2} + \frac{Y_{S_n/2}}{2}\sqrt{S_n} = \frac{2\ln n + X_n\sqrt{2\ln n}}{2} + \frac{Y_{S_n/2}}{2}\sqrt{S_n} \\ &\approx \ln n + \frac{X_n + Y_{\ln n}}{\sqrt{2}}\sqrt{\ln n}; \end{aligned}$$

in the last approximation, we neglect the variations of S_n around $2\ln n$. The Binomial variable is determined by the coin flips ξ_i which are independent of $\mathcal{S}_n(v)$. Thus their (limiting) fluctuations, N and N' , should behave independently. It now follows that

$$(6) \quad \frac{h_n(1) - \ln n}{\sqrt{\ln n}} \approx \frac{1}{\sqrt{2}}(X_n + Y_{\ln n}) \approx \frac{1}{\sqrt{2}}(N + N')$$

where the latter expression has a standard Gaussian distribution. This gives a heuristic of the limiting distribution of $h_n(1)$ without any conditioning.

To prepare for the proof of Theorem 2.5, we next state a lemma describing the joint law of the depth and degree of a given vertex.

Lemma 2.7. *Fix $v \in [n]$, let G be $\text{Geo}(1/2)$ independent of $\mathcal{S}_n(v)$ and let $D = \min\{G, |\mathcal{S}_n(v)|\}$. Then, $d_n(v) \stackrel{\mathcal{L}}{=} D$ and for all $k, l \in \mathbb{N}$,*

$$\mathbf{P}(d_n(v) \geq k, h_n(v) \leq l) = 2^{-k} \mathbf{P}(\text{Bin}(|\mathcal{S}_n(v)| - k, 1/2) \leq l, |\mathcal{S}_n(v)| \geq k).$$

Proof. Any vertex starts as the root of a single-vertex tree. If $|\mathcal{S}_n(v)| = m$, then we flip a fair coin m times and set $d_n(v)$ as the length of the first streak of heads and $h_n(v)$ as the total number of tails; this proves the distributional identity of $d_n(v)$.

Moreover, if $d_n(v) \geq k$, then $|\mathcal{S}_n(v)| \geq k$ and the first k coin flips are determined to be heads, the latter event occurring with probability 2^{-k} . The remaining $|\mathcal{S}_n(v)| - k$ coin flips are independent of the previous tosses. \square

Using Lemma 2.7, we have for all $k \geq m_1 = \lfloor a_1 \ln n \rfloor + b_1$,

$$\mathbf{P}(|\mathcal{S}_n(v)| \geq k | d_n(v) \geq m_1) = \frac{\mathbf{P}(|\mathcal{S}_n(v)| \geq k)}{\mathbf{P}(|\mathcal{S}_n(v)| \geq m_1)} = (1 + o(1))\mathbf{P}(|\mathcal{S}_n(v)| \geq k);$$

the last equality by use the bounds in (4) and the fact that for any $a_1 \in [0, 1]$ and $b_1 \in \mathbb{Z}$, we have $m_1 < (3/2)\ln n$ for n large enough.

Thus, conditioning on the event $\{d_n(1) \geq m_1\}$ does not have a real impact on the distribution of $|\mathcal{S}_n(1)|$. Therefore, $h_n(1)$ depends essentially on $(2 - a_1)\ln n$ fair coin flips. In other words, the conditional law of $h_n(1)$, given that $d_n(1) \geq m_1$ satisfies

$$(7) \quad h_n(1) \approx \text{Bin}(S_n - a\lfloor \log n \rfloor - m, 1/2) \approx (1 - (a_1 \log e)/2)\ln n + \frac{X_n + Y_{\ln n}}{\sqrt{2}}\sqrt{\ln n}.$$

This suggests that, using a suitable choice of renormalizing constants, the conditional law of $h_n(1)$ given that $d_n(1) \geq m_1$ has an asymptotic normal distribution.

To conclude the proof outline for Theorem 2.5, we briefly explain how the depths of distinct vertices are correlated. For $k \geq 2$, the joint distribution of $(h_n(v), v \in [k])$ does not depend only on the sizes of the selection sets $(\mathcal{S}_n(v), v \in [k])$, but also on their overlaps (i.e. on the sets $\mathcal{S}_n(v) \cap \mathcal{S}_n(w)$, for $v, w \in [k]$).

For distinct vertices v, w , let $\lambda_{v,w} = \max\{2 \leq l \leq n : l \in \mathcal{S}_n(v) \cap \mathcal{S}_n(w)\}$. Then, $\lambda_{v,w}$ is the first time that both the trees containing v and w are merged together; moreover, the merging of v, w coincide for the rest of the process. In terms of their depths, this implies that $\kappa_{\lambda_{v,w},v} = 1 - \kappa_{\lambda_{v,w},w}$, i.e. exactly one of v or w increases its depth at step $\lambda_{v,w}$, and also $\kappa_{i,v} = \kappa_{i,w}$, for all $i < \lambda_{v,w}$.

We proceed to outline the contents of the remainder of the paper. In the next section, Section 3, we make rigorous the heuristics in (6) and (7). To do so, we express the cumulative distribution function of $h_n(v)$ as the expected value of a function of $|\mathcal{S}_n|$.

In Section 4, we address the correlations between $(h_n(v), v \in [k])$. We work with the coalescent process stopped at the moment where there are $\ln^2 n$ remaining trees, $F_{\ln^2 n}$. Using $F_{\ln^2 n}$ we define, for each $v \in [n]$, the *truncated selection sets*

$$\mathcal{S}_{n,1}(v) = \mathcal{S}_n(v) \setminus [\ln^2 n]$$

and a partial depth $h_{n,1}(v) = h_{F_{\ln^2 n}}(v)$. In Section 5 we show that if

$$h_n(v) = h_{n,1}(v) + h_{n,2}(v),$$

then $h_{n,2}$ is negligible for the asymptotic distribution of $h_n(v)$; this holds even if we condition on a finite set of vertices (that includes v) to have large degree. Stopping the process, instead, at $F_{\ln \ln n}$ would facilitate the analysis of $h_{n,1}(v) - h_{F_{\ln \ln n}}(v)$, but estimates on $(\mathcal{S}_n(v) \setminus [\ln \ln n], v \in [k])$ would become much more delicate.

In Section 6 we study the joint limiting distribution of $(h_{n,1}(v); v \in [k])$ and complete the proof of Theorem 2.5. Finally, Sections 7 and 8 contain the proofs of Proposition 1.8 and Theorem 1.1, respectively.

3. PROOF OF THEOREM 2.5, CASE $k = 1$

In this section we fix $a \in [0, 1]$, $b \in \mathbb{Z}$ and write $m = m(a, b, n) = \lfloor a \log n \rfloor + b$. We establish the conditional limiting distribution of $(h_n(1) - \mu_a \ln n) / \sqrt{\sigma_a^2 \ln n}$ given that $d_n(1) \geq m$. Our approach consists on averaging over the size of the selection set $\mathcal{S}_n(1)$, and applying the following limit equivalence for the renormalized version of $|\mathcal{S}_n(1)|$.

Lemma 3.1. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a uniformly continuous bounded function, $g : \mathbb{R} \rightarrow \mathbb{R}$ a continuous function and $(g_n : \mathbb{R} \rightarrow \mathbb{R}, n \in \mathbb{N})$ be a sequence of functions uniformly converging to g over any compact set of \mathbb{R} . Let X_n be a sequence of random variables which converges in distribution to X , then*

$$\lim_{n \rightarrow \infty} \mathbf{E}[f(g_n(X_n))] = \mathbf{E}[f(g(X))].$$

Proof of Lemma 3.1. By the Portmanteau theorem we have that

$$\lim_{n \rightarrow \infty} \mathbf{E}[f(g(X_n))] = \mathbf{E}[f(g(X))],$$

so it suffices to prove that

$$\lim_{n \rightarrow \infty} \mathbf{E}[f(g_n(X_n)) - f(g(X_n))] = 0.$$

For an arbitrary $\varepsilon > 0$, let $K = K(\varepsilon) > 0$ be such that $\mathbf{P}(|X| > K) < \varepsilon$ and let $\delta = \delta(\varepsilon)$ such that if $|a - b| < \delta$ then $|f(a) - f(b)| < \varepsilon$. This is possible by the uniform continuity of f . Now, let $n_0 = n_0(\delta, K) \in \mathbb{N}$ be such that $|g_n(x) - g(x)| < \delta$ for all $x \in [-K, K]$ and $n \geq n_0$. This is possible by the uniform convergence of g_n on a bounded set. It follows that

$$|f(g_n(x)) - f(g(x))| < \varepsilon,$$

for any $x \in [-K, K]$ and n large enough. Finally, if $M > 0$ is a bound for f then for n large enough we have

$$\begin{aligned} \mathbf{E}[|f(g_n(X_n)) - f(g(X_n))|] &= \mathbf{E}[|f(g_n(X_n)) - f(g(X_n))| \mathbf{1}_{|X_n| > K}] \\ &\quad + \mathbf{E}[|f(g_n(X_n)) - f(g(X_n))| \mathbf{1}_{|X_n| \leq K}] \\ &\leq 2M \mathbf{P}(|X_n| > K) + \varepsilon \mathbf{P}(|X_n| \leq K) \\ &\leq (2M + 1)\varepsilon. \end{aligned}$$

Since ε is arbitrary, this completes the proof. \square

We describe here a straightforward computation which arises in our proofs.

Lemma 3.2. *Let N be a standard Gaussian variable. Then, for every $x \in \mathbb{R}$ and $b > 0$,*

$$\mathbf{E} \left[\Phi \left(\frac{\sqrt{1+b^2}x - N}{b} \right) \right] = \Phi(x).$$

Proof. Write $a = \sqrt{1+b^2}$, $\sigma = \frac{b}{a}$ and $\mu = \frac{1}{a}$. The expected value can be expressed as a double integral, and changing variables with $y = \frac{aw-z}{b}$ in the second line, we have

$$\begin{aligned} \mathbf{E} \left[\Phi \left(\frac{ax - N}{b} \right) \right] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{ax-z}{b}} \frac{e^{-(y^2+z^2)/2}}{2\pi} dy dz \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^x \frac{\exp\{-\frac{1}{2}[(\frac{aw-z}{b})^2 + z^2]\}}{2\sigma\pi} dw dz \\ &= \int_{-\infty}^x \frac{e^{-w^2/2}}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} \frac{\exp\{-\frac{1}{2\sigma^2}[z - \mu w]^2\}}{\sqrt{2\sigma^2\pi}} dz \right) dw \\ &= \int_{-\infty}^x \frac{e^{-w^2/2}}{\sqrt{2\pi}} dw = \Phi(x). \end{aligned}$$

The third equality holds by the following chain of identities, the key point being that $a^2 = 1 + b^2$,

$$\begin{aligned} \left(\frac{aw - z}{b} \right)^2 + z^2 &= \frac{a^2 w^2}{b^2} - \frac{2awz}{b^2} + \left(\frac{1+b^2}{b^2} \right) z^2 \\ &= w^2 + \frac{a^2}{b^2} \left(\frac{w^2}{a^2} - \frac{2wz}{a} + z^2 \right) \\ &= w^2 + \frac{1}{\sigma^2} \left(z - \frac{w}{a} \right)^2. \end{aligned}$$

□

We first consider the case $m \leq 0$; in other words, the limiting distribution of $h_n(1)$ without conditioning on its degree. For any fixed $x \in \mathbb{R}$, let $G_{n,x} : \mathbb{N} \rightarrow [0, 1]$ be defined as

$$(8) \quad G_{n,x}(t) = \mathbf{P} \left(\text{Bin}(t, 1/2) < x\sqrt{\ln n} + \ln n \right).$$

The motivation behind this definition is that, conditioning on $|\mathcal{S}_n(1)|$ and using (5), we have

$$(9) \quad \mathbf{P} \left(h_n(1) < x\sqrt{\ln n} + \ln n \right) = \mathbf{E} [G_{n,x}(|\mathcal{S}_n(1)|)].$$

The following result describes $G_{n,x}(|\mathcal{S}_n(1)|)$ as a function in terms of $\hat{S}_n = \frac{|\mathcal{S}_n(1)| - 2 \ln n}{\sqrt{2 \ln n}}$ and exploits the Gaussian limit of binomial variables $\text{Bin}(m, p)$ as $m \rightarrow \infty$.

Lemma 3.3. *Let N be a standard Gaussian variable. For any $x \in \mathbb{R}$ fixed,*

$$\lim_{n \rightarrow \infty} \mathbf{E} [G_{n,x}(|\mathcal{S}_n(1)|)] = \mathbf{E} \left[\Phi(\sqrt{2}x - N) \right].$$

Proof. For each $n \in \mathbb{N}$, let $g_{n,x} : \mathbb{R} \rightarrow \mathbb{R}$ be defined as

$$g_{n,x}(r) = (\sqrt{2}x - r) \left(1 + \frac{r}{\sqrt{2 \ln n}} \right)^{-1/2},$$

for $r > -\sqrt{2 \ln n}$, and zero otherwise. Note that $g_{n,x}$ converges to $g_x(r) = \sqrt{2}x - r$, uniformly over bounded intervals as $n \rightarrow \infty$; this is easily proven and we omit the details. Next, we rewrite $\mathbf{E} [G_{n,x}(|\mathcal{S}_n(1)|)]$ as function of \hat{S}_n ; this is to exploit the fact that \hat{S}_n converges in distribution to a standard Gaussian variable by (3). We show that

$$(10) \quad \lim_{n \rightarrow \infty} \mathbf{E} [G_{n,x}(|\mathcal{S}_n(1)|)] = \lim_{n \rightarrow \infty} \mathbf{E} \left[\Phi(g_{n,x}(\hat{S}_n)) \right] = \mathbf{E} \left[\Phi(\sqrt{2}x - N) \right],$$

where N is a standard Gaussian variable. The last equality follows by Lemma 3.1 as the necessary conditions are satisfied: Φ is uniformly continuous and bounded, $g_{n,x}$ converges uniformly over bounded intervals, and \hat{S}_n converges in distribution.

It remains to prove the first equality in (10). Note that

$$G_{n,x}(t) = \mathbf{P} \left(\frac{2\text{Bin}(t, 1/2) - t}{\sqrt{t}} \leq \frac{2x\sqrt{\ln n} + 2\ln n - t}{\sqrt{t}} \right);$$

additionally, letting $t = 2\ln n + r\sqrt{2\ln n}$ we have both $r > -\sqrt{2\ln n}$ and

$$\frac{2x\sqrt{\ln n} + 2\ln n - t}{\sqrt{t}} = \frac{(\sqrt{2}x - r)\sqrt{2\ln n}}{\sqrt{2\ln n - r\sqrt{2\ln n}}} = (\sqrt{2}x - r) \left(\frac{2\ln n - r\sqrt{2\ln n}}{2\ln n} \right)^{-1/2}.$$

For $t \geq 1$, let

$$\mathcal{E}(t) = G_{n,x}(t) - \Phi \left(g_{n,x} \left(\frac{t - 2\ln n}{\sqrt{2\ln n}} \right) \right).$$

By the Berry-Essen theorem for Gaussian approximation, see e.g. [11, Theorem 3.4.9], we have that $|\mathcal{E}(t)| \leq Ct^{-1/2}$ for all $t \geq 1$. Therefore, using the tail bound in (4) for $|\mathcal{S}_n(1)|$, we have as $n \rightarrow \infty$,

$$\mathbf{E} [|\mathcal{E}(|\mathcal{S}_n(1)|)] \leq \mathbf{E} [|\mathcal{E}(|\mathcal{S}_n(1)|)] \leq \mathbf{P} (|\mathcal{S}_n(1)| \leq \ln n) + C(\ln n)^{-1/2} \rightarrow 0.$$

This completes the proof as (10) follows from

$$\lim_{n \rightarrow \infty} \mathbf{E} [G_{n,x}(|\mathcal{S}_n(1)|)] = \lim_{n \rightarrow \infty} \mathbf{E} \left[\Phi(g_{n,x}(\hat{S}_n)) \right] + \lim_{n \rightarrow \infty} \mathbf{E} [\mathcal{E}(|\mathcal{S}_n(1)|)],$$

where both limits in the right-hand side exist and the last one vanishes. \square

Despite Lemma 3.4 below being a stronger statement than Lemma 3.3, we decided to present the detailed proof of Lemma 3.3 as the computations are easier to follow. In particular, Lemmas 3.2 and 3.3 together imply that for any $x \in \mathbb{R}$,

$$(11) \quad \lim_{n \rightarrow \infty} \mathbf{P} \left(h_n(1) < x\sqrt{\ln n} + \ln n \right) = \Phi(x),$$

which formalizes the heuristic in (6) and already yields a particular case of Theorem 2.5.

We now proceed to deal with the case $k = 1$ and a non-trivial conditioning in Theorem 2.5. For any $d, l \in \mathbb{N}$ let $\tilde{G}_{d,l} : \mathbb{N} \rightarrow [0, 1]$ be defined as

$$(12) \quad \tilde{G}_{d,l}(t) = \mathbf{P} (\text{Bin}(t - d, 1/2) < l) \mathbf{1}_{[t \geq d]}.$$

By Lemma 2.7 we get

$$(13) \quad \begin{aligned} \mathbf{P} (h_n(1) \leq l, d_n(1) \geq d) &= 2^{-d} \sum_{t \geq d} \mathbf{P} (h_n(1) \leq l \mid |\mathcal{S}_n(1)| = t) \mathbf{P} (|\mathcal{S}_n(1)| = t) \\ &= 2^{-d} \mathbf{E} [\tilde{G}_{d,l}(|\mathcal{S}_n(1)|)]. \end{aligned}$$

Recall the next definitions given for Theorem 1.3; for $a \in [0, 1]$, let $\mu_a = 1 - (a \log e)/2$ and $\sigma_a^2 = 1 - (a \log e)/4$.

Lemma 3.4. *Fix $a \in [0, 1]$, $b \in \mathbb{Z}$ and let $x \in \mathbb{R}$. Write $m = m(a, b, n) = \lfloor a \log n \rfloor + b$ and $l = l(a, x, n) = x\sqrt{\sigma_a^2 \ln n} + \mu_a \ln n$. If $m \geq 0$ then*

$$\mathbf{E} [\tilde{G}_{m,l}(|\mathcal{S}_n(1)|)] = \mathbf{E} \left[\Phi \left(\frac{\sqrt{2\sigma_a^2}x - N}{\sqrt{\mu_a}} \right) \right].$$

Proof. The proof uses Lemma 3.1 and follows the same approach as in Lemma 3.3. We also use the renormalization \hat{S}_n . Fix a, b, x and set m, l as given in the statement. For the rest of the proof, write $\mu = \mu_a$ and $\sigma = \sigma_a$. We show that

$$(14) \quad \lim_{n \rightarrow \infty} \mathbf{E} \left[\tilde{G}_{m,l}(|\mathcal{S}_n(1)|) \right] = \lim_{n \rightarrow \infty} \mathbf{E} \left[\Phi(\tilde{g}_{n,a,x}(\hat{S}_n)) \right],$$

where $\tilde{g}_{n,a,x} : \mathbb{R} \rightarrow \mathbb{R}$ are functions, defined below, such that $\tilde{g}_{n,a,x}(r)$ converges to $\tilde{g}_{a,x}(r) = \frac{\sqrt{2}\sigma x - t}{\sqrt{\mu}}$, uniformly over bounded sets, as $n \rightarrow \infty$. Once (14) is established, the result follows by Lemma 3.1. To do so, we are required to bound the error of approximating $\tilde{G}_{m,l}(|\mathcal{S}_n(1)|)$ with $\Phi(\tilde{g}_{n,a,x}(\hat{S}_n))$.

Now, write $\varepsilon = \varepsilon(a, n) = a \log n - \lfloor a \log n \rfloor$; then $m = \lfloor a \log n \rfloor + b = 2(1 - \mu) \ln n + b - \varepsilon$. A direct calculation shows that

$$\tilde{g}_{n,a,x}(r) = \frac{\sqrt{2}\sigma x - r}{\sqrt{\mu}} \left(1 + \frac{r\sqrt{2 \ln n} - b + \varepsilon}{2\mu \ln n} \right)^{-1/2} + \left(\frac{2\mu \ln n + r\sqrt{2 \ln n} - b + \varepsilon}{(b - \varepsilon)^2} \right)^{-1/2}$$

if $r \geq \frac{-2\mu \ln n + b - \varepsilon}{\sqrt{2 \ln n}}$, and zero otherwise. The uniform convergence of $\tilde{g}_{n,a,x}$ is straightforward, but we omit the details. For $t \geq 1$, let

$$\mathcal{E}(t - m) = \tilde{G}_{m,l}(t) - \Phi \left(\tilde{g}_{n,a,x} \left(\frac{t - 2 \ln n}{\sqrt{2 \ln n}} \right) \right).$$

By the Berry-Essen theorem, see e.g. [11, Theorem 3.4.9], we have that $|\mathcal{E}(t)| \leq Ct^{-1/2}$. Finally, for n large enough, $m < (3/2) \ln n$ and so, having $|\mathcal{S}_n(1)| > (7/4) \ln n$ implies $|\mathcal{S}_n(1)| - m > (1/4) \ln n$. By (4) we get,

$$\mathbf{E}[\mathcal{E}(|\mathcal{S}_n(1)| - m)] \leq \mathbf{E}[|\mathcal{E}(|\mathcal{S}_n(1)| - m)|] \leq \mathbf{P}(|\mathcal{S}_n(1)| \leq (2 - 1/4) \ln n) + 2C(\ln n)^{-1/2} = o(1).$$

This completes the proof as

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\tilde{G}_{m,l}(|\mathcal{S}_n(1)|) \right] = \lim_{n \rightarrow \infty} \mathbf{E} \left[\Phi(\tilde{g}_{n,a,x}(\hat{S}_n)) \right] = \lim_{n \rightarrow \infty} \mathbf{E}[\mathcal{E}(|\mathcal{S}_n(1)| - m)],$$

and the last limit vanishes. \square

Proof of Theorem 2.5, case $k = 1$. Fix $a \in [0, 1]$, $b \in \mathbb{Z}$ and $x \in \mathbb{R}$. Let $m = m(a, b, n) = \lfloor a \log n \rfloor + b$ and $l = l(a, x, n) = x\sqrt{\sigma_a^2 \ln n} + \mu_a \ln n$. Our goal is to show that

$$\lim_{n \rightarrow \infty} \mathbf{P}(h_n(1) < l \mid d_n(1) \geq m) = \Phi(x).$$

If $m \leq 0$ then $a = 0$. The result then follows by (11) since $\mu_a = \sigma_a = 1$, and so

$$\mathbf{P}(h_n(1) < l \mid d_n(1) \geq m) = \mathbf{P}(h_n(1) < x\sqrt{\ln n} + \ln n).$$

Consider now the case $m > 0$. Note that $m = \lfloor a \log n \rfloor + b \leq \frac{3}{2} \ln n$ for n large enough. Therefore, by Lemma 2.7 and (4), we have

$$(15) \quad \lim_{n \rightarrow \infty} 2^m \mathbf{P}(d_n(1) \geq m) = \lim_{n \rightarrow \infty} \mathbf{P}(|\mathcal{S}_n(1)| \geq m) = 1.$$

Using the equations (13), (15), and Lemma 3.4 we get that for any $x \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \mathbf{P}(h_n(1) < l \mid d_n(1) \geq m) = \lim_{n \rightarrow \infty} 2^{-m} \frac{\mathbf{E}[\tilde{G}_{m,l}(|\mathcal{S}_n(1)|)]}{\mathbf{P}(|\mathcal{S}_n(1)| \geq m)} = \mathbf{E} \left[\Phi \left(\frac{\sqrt{2}\sigma_a x - N}{\sqrt{\mu_a}} \right) \right].$$

We use the fact that $2\sigma_a^2 = 1 + \mu_a$ to apply Lemma 3.2 to the last term above. This yields the desired result. \square

In the following section we lay down the necessary approximations to obtain a generalization of (13) to several vertices.

4. TRUNCATED SELECTION SETS

In this section we fix $k \geq 2$ and consider the depths of vertices in $F_{\ln^2 n}$. Recall from Section 2.2 that the truncated selection sets are defined by $\mathcal{S}_{n,1}(v) = \mathcal{S}_n(v) \setminus [\ln^2 n]$ for $v \in [n]$. Let $\Omega_1 = \mathcal{P}(\{\log^2 n + 1, \dots, n\})$. For the remainder of the paper we use, e.g. $\bar{\mathcal{S}}_{n,1} \in \Omega_1^k$ to denote the vector $(\mathcal{S}_{n,1}(i), i \in [k])$.

Our main objective is showing that $(\mathcal{S}_{n,1}(i), i \in [k])$ behave, asymptotically, as if they were independent sets; see Proposition 4.6. Also, the conditional law of the depths $(h_{n,1}(i), i \in [k])$ given the truncated selection sets $(\mathcal{S}_{n,1}(i), i \in [k])$ can be approximated by the law of k independent Binomial variables. This holds even if we condition on the final degrees $(d_n(i), i \in [k])$; see Proposition 4.2 and Remark 4.3. These properties are crucial to establishing Theorem 2.5 in full generality.

The choice of halting the process at $F_{\ln^2 n}$, and not e.g. $F_{\ln \ln n}$, implies that we must also provide the limiting distribution of $(h_n(i) - h_{n,1}(i))/\sqrt{\ln n}$ (a priori $h_n(i) - h_{n,1}(i) \leq \ln^2 n$). Nevertheless, we use $F_{\ln^2 n}$ since it allows to use simple arguments in the estimates of Proposition 4.6 below.

Note that, $\mathbf{E}[|\mathcal{S}_{n,1}(v)|] = 2 \ln n - 2 \ln \ln n + o(1) = 2 \ln n(1 + o(1))$ and thus, similar to (4), we get concentration of $|\mathcal{S}_{n,1}(v)|$ around $2 \ln n$ and a normal asymptotic limit.

Fact 4.1. *For any $v \in [n]$ and $\varepsilon > 0$, $\mathbf{P}(|\mathcal{S}_{n,1}(v)| - 2 \ln n| > \varepsilon \ln n) = o(1)$ and*

$$\frac{|\mathcal{S}_{n,1}(v) - 2 \ln n|}{\sqrt{2 \ln n}} \xrightarrow{\mathcal{L}} N;$$

where N is a standard Gaussian variable.

The following proposition is used to obtain independent limiting distributions for the depths of k vertices in the final tree $T^{(n)}$.

Proposition 4.2. *Fix $\bar{m}, \bar{l} \in \mathbb{N}^k$. For all $\bar{J} \in \Omega_1^k$ such that $\{J_i, i \in [k]\}$ are pairwise disjoint, we have*

$$\mathbf{P}(h_{n,1}(i) \leq l_i, i \in [k] | \bar{\mathcal{S}}_{n,1} = \bar{J}) = \prod_{i=1}^k \mathbf{P}(\text{Bin}(|J_i|, 1/2) \leq l_i),$$

$$\mathbf{P}(d_{F_{\ln^2 n}}(i) \geq m_i, h_{n,1}(i) \leq l_i, i \in [k] | \bar{\mathcal{S}}_{n,1} = \bar{J}) = 2^{-\sum m_i} \prod_{i=1}^k \mathbf{P}(\text{Bin}(|J_i| - m_i, 1/2) \leq l_i) \mathbf{1}_{|J_i| \geq m_i}.$$

Proof. Fix $\bar{m}, \bar{l} \in \mathbb{N}^k$ and $\bar{J} \in \Omega_1^k$ as given in the statement. Once the sets $(\mathcal{S}_{n,1}(i), i \in [k])$ are fixed, the depth $h_{n,1}(i)$ of $i \in [k]$ in $F_{\ln^2 n}$ is determined by the variables $(\xi_j, j \in \mathcal{S}_{n,1}(i))$. Consequently, given that $\bar{\mathcal{S}}_{n,1} = \bar{J}$ the conditional law of the degrees and depths of vertices in $F_{\ln^2 n}$ depend on disjoint sets of independent variables. Therefore, we can decouple the event $\{d_{F_{\ln^2 n}}(i) \geq m_i, h_{n,1}(i) \leq l_i\}$.

The first equality in the statement corresponds to the case when $m_i = 0$ for all $i \in [k]$. Now, for the second equality we first note that the product of indicator functions follows since $d_{F_{\ln^2 n}}(i) \geq d_i$ for all $i \in [k]$ occurs only if $|\mathcal{S}_{n,1}(i)| \geq m_i$ for each $i \in [k]$. Then, for each $i \in [k]$ we flip $|\mathcal{S}_{n,1}(i)|$ independent fair coins. The first m_i coins must be heads and this occurs with probability 2^{-m_i} . The number of tails in the remaining coin flips determine the depth $h_{n,1}(i)$; this is distributed as $\text{Bin}(|\mathcal{S}_{n,1}(i)| - m_i, 1/2)$. \square

Remark 4.3. *Furthermore, if $\bar{J} \in \Omega_1^k$ is such that $|J_i| \geq m_i$ for all $i \in [k]$, then*

$$\{d_n(i) \geq m_i, i \in [k], \bar{\mathcal{S}}_{n,1} = \bar{J}\} = \{d_{F_{\ln^2 n}}(i) \geq m_i, i \in [k], \bar{\mathcal{S}}_{n,1} = \bar{J}\}.$$

Now, with high-probability, vertices in $[k]$ still belong to distinct trees in $F_{\ln^2 n}$ which implies that the truncated selection sets $(\mathcal{S}_{n,1}(i), i \in [k])$ are disjoint. To see this, let us define

$$\tau_k = \max\{2 \leq j \leq n : s_{j,v} = s_{j,w} = 1 \text{ for some distinct } v, w \in [k]\}.$$

Recall that the trees in F_j are ordered in increasing order of their least element. By definition of τ_k , $|\{a_j, b_j\} \cap [k]| \leq 1$ for $j > \tau_k$. Thus, at no point $j \geq \tau_k$ are $T_j(v)$ and $T_j(w)$ merged, for distinct

$v, w \in [k]$. In other words, $T_j(i) = i$ for all $i \in [k]$, $j \geq \tau_k$. Therefore,

$$(16) \quad \{\tau \leq \ln^2 n\} = \{(\mathcal{S}_{n,1}(i), i \in [k]) \text{ are pairwise disjoint}\}.$$

Fact 4.4. Fix an integer $k \geq 2$. For n large enough,

$$\mathbf{P}(\tau_k > \ln^2 n) \leq 2k^2 \ln^{-2} n.$$

Proof. By definition, $T_j(i) = T_i^{(j)}$ for all $i \in [k]$ and $j \geq \tau_k$. Therefore,

$$\mathbf{P}(\tau_k \leq l) = \prod_{j=l+1}^n \mathbf{P}(|\{a_j, b_j\} \cap [k]| < 2) = \prod_{j=l+1}^n \left(1 - \frac{k(k-1)}{j(j-1)}\right) \geq \prod_{j=l}^{\infty} \left(1 - \frac{k^2}{j^2}\right).$$

The second equality is since the pairs $(\{a_j, b_j\}, 2 \leq j \leq n)$ are chosen independently and uniformly at random. For the next approximation we use that $1 - x > e^{-2x}$ for $x > 0$ sufficiently small and that $e^{-x} > 1 - x$ for all $x \in \mathbb{R}$. Then, letting $l = \ln^2 n$ and n large enough, we have

$$\prod_{j=\ln^2 n}^{\infty} \left(1 - \frac{k^2}{j^2}\right) > 1 - \sum_{j=\ln^2 n}^{\infty} \frac{2k^2}{j^2} > 1 - 2k^2 \int_{\ln^2 n}^{\infty} x^{-2} dx = 1 - 2k^2 \ln^{-2} n. \quad \square$$

Finally, we consider the following family of sets as representing the bulk of the probability measure induced by k truncated sets. We add the parameter $\delta > 0$ to cover the distinct possible values of $\bar{a} \in [0, 1]^k$ in Theorem 2.5. For $\delta \in (0, 2)$ let

$$(17) \quad \mathcal{B}_{n,k,\delta} = \{\bar{J} \in \Omega_1^k : (J_1, \dots, J_k) \text{ are pairwise disjoint and } ||J_i| - 2 \ln n| \leq \delta \ln n, i \in [k]\}.$$

Lemma 4.5. Fix an integer $k \geq 2$ and $\delta \in (0, 2)$. Then

$$\mathbf{P}(\bar{\mathcal{S}}_{n,1} \in \mathcal{B}_{n,k,\delta}) = 1 + o(1).$$

Proof. This follows directly from (16) and Facts 4.1 and 4.4;

$$\mathbf{P}(\bar{\mathcal{S}}_{n,1} \notin \mathcal{B}_{n,k,\delta}) \leq \mathbf{P}(\tau_k \geq \ln^2 n) + k\mathbf{P}(|\mathcal{S}_{n,1}(i)| - 2 \ln n| < \delta \ln n) = o(1). \quad \square$$

Let $(\mathcal{R}_n(i), i \in [k])$ be k independent copies of $\mathcal{S}_{n,1}(1)$. We use sets $\mathcal{B}_{n,k,\delta}$ to make explicit the claim that $(\mathcal{S}_{n,1}(i), i \in [k])$ are asymptotically independent; this occurs uniformly on such $\mathcal{B}_{n,k,\delta}$.

Proposition 4.6. Fix an integer $k \geq 2$ and $\delta \in (0, 2)$. Uniformly for $\bar{J} \in \mathcal{B}_{n,k,\delta}$,

$$\mathbf{P}(\bar{\mathcal{S}}_{n,1} = \bar{J}) = (1 + o(1))\mathbf{P}(\bar{\mathcal{R}}_n = \bar{J}).$$

The remainder of the section is devoted to proving Proposition 4.6, and to do so we fix $\delta \in (0, 2)$ and $\bar{J} \in \mathcal{B}_{n,k,\delta}$. The notation we define below does not reflect the dependency on \bar{J} . We use the index m with $\ln^2 n < m \leq n$ unless otherwise specified. Recall that $\mathcal{S}_{n,1}(i) = \{m : s_{m,i} = 1\}$. Similarly, for all m and $i \in [k]$, let $r_{m,i}$ be the random indicator of $m \in \mathcal{R}_n(i)$ and let $j_{m,i} = \mathbf{1}_{[m \in J_i]}$. Also, let $\sigma_m = \sum_{i \in [k]} j_{m,i}$ and note that from the choice of \bar{J} we have that $\sigma_m \leq 1$ for all m .

Claim 4.7. For each m , let $A_m = \{s_{m,i} = j_{m,i}, i \in [k]\}$. Then

$$(18) \quad \mathbf{P}(A_m | A_l, m < l \leq n) = \begin{cases} \frac{(m-k)(m-k-1)}{m(m-1)} & \text{if } \sigma_m = 0, \\ \frac{2(m-k)}{m(m-1)} & \text{if } \sigma_m = 1. \end{cases}$$

and furthermore,

$$\mathbf{P}(\mathcal{S}_{n,1}(i) = J_i, i \in [k]) = \prod_m \mathbf{P}(A_m | A_l, m < l \leq n).$$

Proof. The second equality follows since $\{\mathcal{S}_{n,1}(i) = J_i; i \in [k]\} = \{\cap_m A_m\}$. We proceed to prove (18) by induction on $n - m$. For $m = n$, the formula is trivial. For $m < n$ note that the condition $\{A_l, m < l \leq n\}$ implies that $\sigma_l \leq 1$ for all $m < l \leq n$. That is, there has been no merges of distinct trees $T_l(v), T_l(w)$ for $v, w \in [k]$. In particular, $T_m(i) = T_i^{(m)}$ for all $i \in [k]$. If $\sigma_m = 0$, then none of these trees are selected to be merged in the next step, and this occurs with probability $\frac{(m-k)(m-k-1)}{m(m-1)}$. If $\sigma_m = 1$, then there is exactly one vertex $i \in [k]$ which is selected and the other tree is selected among $(m - k)$ trees. \square

Similarly, we have the following estimates for $(\mathcal{R}_n(i), i \in [k])$.

Claim 4.8. *For each m , let $A'_m = \{r_{m,i} = j_{m,i}, i \in [k]\}$. Then*

$$(19) \quad \mathbf{P}(A'_m) = \begin{cases} \left(1 - \frac{2}{m}\right)^k & \text{if } \sigma_m = 0, \\ \frac{2}{m} \left(1 - \frac{2}{m}\right)^{k-1} & \text{if } \sigma_m = 1. \end{cases}$$

and furthermore,

$$\mathbf{P}(\mathcal{R}_n(i) = J_i, i \in [k]) = \prod_m \mathbf{P}(A_m).$$

Proof. It is clear that $\{\mathcal{R}_n(i) = J_i; i \in [k]\} = \{\cap_m A'_m\}$. Observe that the events A'_m are independent. Also, (19) follows immediately from the distribution of $(r_{m,i}, i \in [k])$ and the fact that these variables are independent. \square

Proposition 4.6 is obtained by comparing the two products in the claims above. The following claim relates each of the terms in (18) and (19). Let

$$\begin{aligned} p_{m,0} &= \frac{(m-k)(m-k-1)}{m(m-1)}, & q_{m,0} &= \left(1 - \frac{2}{m}\right)^k, \\ p_{m,1} &= \frac{2(m-k)}{m(m-1)}, & q_{m,1} &= \frac{2}{m} \left(1 - \frac{2}{m}\right)^{k-1}. \end{aligned}$$

Claim 4.9. *There exists a constant $c = c(k) > 0$ such that for m large enough, we have*

$$\begin{aligned} q_{m,0} &> p_{m,0} > q_{m,0} \left(1 - \frac{c}{m^2}\right), \\ q_{m,1} &< p_{m,1} < q_{m,1} \left(1 + \frac{c}{m}\right). \end{aligned}$$

Proof. First we prove the bounds on $p_{m,0}$. Note that $p_{m,0} = 1 - \frac{2k}{m} + \frac{k(k-1)}{m(m-1)}$ and so

$$(20) \quad 0 < q_{m,0} - p_{m,0} = -\frac{k(k-1)}{m(m-1)} + \frac{2k(k-1)}{m^2} + O(m^{-3}) = O(m^{-2}).$$

The upper bound on $p_{m,0}$ follows from the first inequality in (20). For the lower bound, use that $q_{m,0} \rightarrow 1$ as $m \rightarrow \infty$ then

$$\frac{q_{m,0} - p_{m,0}}{q_{m,0}} = \frac{q_{m,0} - p_{m,0}}{1 + o(1)} = O(m^{-2}).$$

The bounds on $p_{m,1} = \frac{2}{m} - \frac{2(k-1)}{m(m-1)}$ are obtained similarly. We use that $mq_{m,1} \rightarrow 2$ as $m \rightarrow \infty$ and

$$0 < p_{m,1} - q_{m,1} = -\frac{2(k-1)}{m(m-1)} + \frac{4(k-1)}{m^2} + O(m^{-3}) = O(m^{-2}). \quad \square$$

Proof of Proposition 4.6. Fix $\delta \in (0, 2)$ and $k \geq 2$. The bounds we give below do not depend on the choice of $\bar{J} \in \mathcal{B}_{n,k,\delta}$ and so the bounds obtained are uniform in $\mathcal{B}_{n,k,\delta}$. By Claims 4.7 and 4.8, it suffices to prove that

$$\prod_m \mathbf{P}(A_m | A_l, m < l \leq n) = (1 + o(1)) \prod_m \mathbf{P}(A'_m).$$

The lower bounds in Claim 4.9 give, for m large enough,

$$\begin{aligned}
\mathbf{P}(A_m, \ln^2 n < m \leq n) &= \prod_{m=\ln^2 n}^n \mathbf{P}(A_m | A_l, m < l \leq n) \\
&\geq \prod_{m=\ln^2 n}^n \mathbf{P}(A'_m) \left(1 - \frac{c}{m^2}\right) \\
&= \mathbf{P}(A'_m, \ln^2 n < m \leq n) \prod_{m=\ln^2 n}^n \left(1 - \frac{c}{m^2}\right) \\
&\geq \mathbf{P}(A'_m, \ln^2 n < m \leq n) \left(1 - \frac{2c}{\ln^2 n}\right).
\end{aligned}$$

The last equality follows in the same manner as the bound for $\mathbf{P}(\tau_k \leq \ln^2 n)$ obtained in Fact 4.5. Now, using the upper bounds in Claim 4.9 we have, for m large enough,

$$\begin{aligned}
\mathbf{P}(A_m, \ln^2 n < m \leq n) &= \prod_{m=\ln^2 n}^n \mathbf{P}(A_m | A_l, m < l \leq n) \\
&\leq \prod_{m=\ln^2 n}^n \mathbf{P}(A'_m) \left(1 + \frac{c}{m} \mathbf{1}_{[\sigma_m=1]}\right) \\
&= \mathbf{P}(A'_m, \ln^2 n < m \leq n) \prod_{m:\sigma_m=1} \left(1 + \frac{c}{m}\right) \\
&\leq \mathbf{P}(A'_m, \ln^2 n < m \leq n) \left(1 + \frac{2(2+\delta)ck}{\ln n}\right).
\end{aligned}$$

In the last inequality we use that $\sum_m \sigma_m \leq (2+\delta)k \ln n$ by the second condition on $\mathcal{B}_{n,k,\delta}$. Thus,

$$\prod_{\sigma_m=1} \left(1 + \frac{c}{m}\right) \leq \exp\left(\frac{(2+\delta)ck \ln n}{\ln^2 n}\right) < 1 + \frac{(2+\delta)2ck}{\ln n}.$$

In the first inequality we use that $m \geq \ln^2 n$ and $1+x \leq e^x$ for all $x \in \mathbb{R}$; for the second inequality, we use that $e^x < 1+2x$ for x sufficiently small. \square

5. NEGLIGIBLE DEPTH INCREASE

In this section we fix $k \geq 2$ and prove that the main contribution to $(h_n(i), i \in [k])$ is already found in $F_{\ln^2 n}$. Recall that $h_{n,1}(i) = h_{F_{\ln^2 n}}(i)$ and $h_{n,2}(i) = h_n(i) - h_{n,1}(i)$, for $i \in [n]$. The key observation in this section is that the coalescence after $F_{\ln^2 n}$ can be compared with an independent $\ln^2 n$ -coalescent.

Fact 5.1. *For any $v \in [n]$, $h_{n,2}(v)$ is stochastically dominated by $|\mathcal{S}_{\ln^2 n}(v)|$.*

Proof. In an n -coalescent $\mathbf{C} = (F_n, \dots, F_1)$ we have that $h_{F_i}(v) = \sum_{j=i+1}^n h_{j,v}$ with $h_{j,v} \leq s_{j,v}$. Thus,

$$h_{n,2} = h_n(v) - h_{n,1}(v) = \sum_{j=2}^{\ln^2 n} h_{j,v} \leq \sum_{j=2}^{\ln^2 n} s_{j,v}.$$

The result then follows since for any $m \leq n$, we have that $|\mathcal{S}_m(v)| \stackrel{\mathcal{L}}{=} \sum_{j=2}^m s_{j,v}$. \square

Lemma 5.2. *For any vertex $i \in [n]$, we have $\frac{h_{n,2}(i)}{\sqrt{\ln n}} \rightarrow 0$, in probability as $n \rightarrow \infty$.*

Proof. By Fact 5.1, it suffices to prove that for every $\varepsilon > 0$,

$$(21) \quad \mathbf{P}\left(|\mathcal{S}_{\ln^2 n}(i)| > \varepsilon \sqrt{\ln n}\right) = o(1).$$

Write $m = \ln^2 n$ and note that $\mathbf{E}[|\mathcal{S}_m(i)|] = 2 \ln \ln n + O(1)$ and so $\delta = \frac{\varepsilon \sqrt{\ln n} - \mathbf{E}[|\mathcal{S}_m(i)|]}{\mathbf{E}[|\mathcal{S}_m(i)|]} > 0$ for n large enough. Therefore, Lemma 4.1 yields

$$\mathbf{P}\left(|\mathcal{S}_{\ln^2 n}(i)| > \varepsilon \sqrt{\ln n}\right) \leq \mathbf{P}\left(|\mathcal{S}_m(i)| - \mathbf{E}[|\mathcal{S}_m(i)|] > \delta \mathbf{E}[|\mathcal{S}_m(i)|]\right) = o(1). \quad \square$$

In fact, for $i \in [k]$, $h_{n,2}(i)$ is also negligible when we condition on the vertices in $[k]$ to have large degree. Let $\Omega = \mathcal{P}([n])$ be the power set of $[n]$ and fix $\bar{m} \in \mathbb{N}^k$. Let

$$\begin{aligned} \mathcal{A}_{\bar{m}} &= \{\bar{J} \in \Omega^k : \mathbf{P}(\bar{\mathcal{S}}_n = \bar{J}, d_n(i) \geq m_i, i \in [k]) > 0\}, \\ \mathcal{L}_{\bar{m}} &= \{\bar{J} \in \Omega^k : |J_i \setminus [\ln^2 n]| \geq m_i, i \in [k]\}. \end{aligned}$$

Lemma 5.3. *Fix $\bar{m} \in \mathbb{N}^k$. For any $s \in \mathbb{N}$ and $i \in [k]$, if $\bar{J} \in \mathcal{A}_{\bar{m}}$ we have*

$$\mathbf{P}(h_{n,2}(i) \geq s, d_n(j) \geq m_j, j \in [k] | \bar{\mathcal{S}}_n = \bar{J}) \leq 2^{-\sum_j m_j};$$

if $\bar{J} \in \mathcal{A}_{\bar{m}} \cap \mathcal{L}_{\bar{m}}$, then

$$\mathbf{P}(h_{n,2}(i) \geq s, d_n(j) \geq m_j, j \in [k] | \bar{\mathcal{S}}_n = \bar{J}) \leq 2^{-\sum_j m_j} \mathbf{P}(\mathcal{S}_{\ln^2 n}(i) \geq s).$$

Proof. Recall that the degree of a vertex $i \in [k]$ is determined by the first streak of selection times $j \in \mathcal{S}_n(i)$ where $h_{j,i} = 0$. If $\bar{J} \in \mathcal{A}_{\bar{m}}$, then the event $\bar{\mathcal{S}}_n = \bar{J}$ has the property that the set of the first m_i selection times in $\mathcal{S}_n(i)$ are pairwise disjoint for all $i \in [k]$; otherwise $\mathbf{P}(\bar{\mathcal{S}}_n = \bar{J}, d_n(i) \geq m_i, i \in [k]) = 0$. It then follows that

$$\mathbf{P}(d_n(j) \geq m_j, j \in [k] | \bar{\mathcal{S}}_n = \bar{J}) = 2^{-\sum_j m_j},$$

which yields the first inequality. For the second inequality it remains to prove that for $\bar{J} \in \mathcal{A}_{\bar{m}} \cap \mathcal{L}_{\bar{m}}$,

$$\mathbf{P}(h_{n,2}(i) \geq s | d_n(j) \geq m_j, j \in [k], \bar{\mathcal{S}}_n = \bar{J}) \leq \mathbf{P}(\mathcal{S}_{\ln^2 n}(i) \geq s).$$

In this case, the event $\{d_n(j) \geq m_j, j \in [k]\}$ is already determined by the forest $F_{\ln^2 n}$. Consequently, the remaining selection times $\mathcal{S}_n(i) \cap [\ln^2 n]$, which determine $h_{n,2}(i)$, are independent of the conditioning event and so the argument in Fact 5.1 can be applied. \square

The next lemma uses a result from [2], whose proof can be derived from this work but we omit its proof for brevity.

Proposition 5.4 (Proposition 4.2 in [2]). *Fix $c \in (0, 2)$ and $k \in \mathbb{N}$. There exists $\beta = \beta(c, k) > 0$ such that uniformly over positive integers $m_1, \dots, m_k < c \ln n$,*

$$\mathbf{P}(d_n(j) \geq m_j, j \in [k]) = 2^{-\sum_j m_j} (1 + o(n^{-\beta})).$$

Lemma 5.5. *Fix $c \in (0, 2)$. If $m_j = m_j(n) < c \ln n$ for all $j \in [k]$, then for any $i \in [k]$,*

$$\mathbf{P}(h_{n,2}(i) \geq \varepsilon \sqrt{\ln n}, d_n(j) \geq m_j, j \in [k]) \rightarrow 0.$$

Proof. Let \bar{m} satisfy the conditions of the statement. By Lemma 5.3, we have for any $i \in [k]$ and $s \in \mathbb{N}$,

$$\begin{aligned} \mathbf{P}(h_{n,2}(i) \geq s, d_n(j) \geq m_j, j \in [k]) &= \sum_{\bar{J} \in \mathcal{A}_{\bar{m}}} \mathbf{P}(\bar{\mathcal{S}}_n = \bar{J}, h_{n,2}(i) \geq s, d_n(j) \geq m_j, j \in [k]) \\ &\leq 2^{-d} [\mathbf{P}(|\mathcal{S}_{\ln^2 n}(i)| \geq s) + \mathbf{P}(\bar{\mathcal{S}}_n \notin \mathcal{L}_{\bar{m}})]. \end{aligned}$$

If $s = \varepsilon \sqrt{\ln n}$, the first probability in the last line vanishes as $n \rightarrow \infty$ by (21). Also, by (4) we get

$$\mathbf{P}(\bar{\mathcal{S}}_n \notin \mathcal{L}_{\bar{m}}) \leq k \mathbf{P}(|\mathcal{S}_n(1)| < c \ln n) = o(1).$$

Therefore

$$\mathbf{P}(h_{n,2}(i) \geq \varepsilon \sqrt{\ln n} | d_n(j) \geq m_j, j \in [k]) = \frac{2^{-\sum_j m_j} o(1)}{\mathbf{P}(d_n(j) \geq m_j, j \in [k])};$$

and the proof is completed since $\mathbf{P}(d_n(j) \geq m_j, j \in [k]) = 2^{-\sum_j m_j} (1 + o(1))$ by Proposition 5.4. \square

6. PROOF OF THEOREM 2.5, CASE $k \geq 2$

Fix an integer $k \geq 2$. We would like to express $\mathbf{P}\left(h_n(i) < x_i \sqrt{\ln n} + \ln n, i \in [k]\right)$ as a product of expectations of the form in (9) since this would yield the independence of the limiting variables. However we have seen previously that this is not possible largely due to the correlations between the selection sets of vertices $i \in [k]$. Instead we consider the depths $(h_{n,1}(i), i \in [k])$ and exploit the fact that $(\mathcal{S}_{n,1}(v), v \in [k])$ are asymptotically independent. Given a measure μ , we write \mathbf{E}_μ for expectations with respect to μ .

Lemma 6.1. *For each $n \in \mathbb{N}$, let μ_n and ν_n be probability measures in a space Ω_n . Let $B_n \subset \Omega_n$ be such that, uniformly for each $\omega \in B_n$, $\mu_n(\omega) = (1 + o(1))\nu_n(\omega)$; and $\mu_n(B_n) = 1 + o(1) = \nu_n(B_n)$.*

If $f_n, g_n \in \Omega_n \rightarrow \mathbb{R}$ are bounded and $f_n(\omega) = g_n(\omega)$ for all $\omega \in B_n$; then

$$\mathbf{E}_{\mu_n}[f_n] = (1 + o(1))\mathbf{E}_{\nu_n}[g_n] + o(1).$$

Proof. In the subspace B_n we can interchange f_n and g_n ; and since the approximation $\mu_n(\omega) = (1 + o(1))\nu_n(\omega)$ is uniform over $\omega \in B_n$ we have that

$$\mathbf{E}_{\mu_n}[f_n \mathbf{1}_{B_n}] = (1 + o(1))\mathbf{E}_{\nu_n}[f_n \mathbf{1}_{B_n}] = (1 + o(1))\mathbf{E}_{\nu_n}[g_n \mathbf{1}_{B_n}].$$

The result follows by noting that

$$\mathbf{E}_{\mu_n}[f_n \mathbf{1}_{\Omega_n \setminus B_n}] - (1 + o(1))\mathbf{E}_{\mu_n}[g_n \mathbf{1}_{\Omega_n \setminus B_n}] = o(1);$$

which is a straightforward consequence of f_n and g_n being bounded and that the measure of $\Omega_n \setminus B_n$ vanishes for both measures as $n \rightarrow \infty$. \square

Similar to Section 3 above, we will start with the unconditional case; that is, the limiting distribution of $(h_{n,1}(i), i \in [k])$.

Proposition 6.2. *Fix an integer $k \geq 2$. For any $\bar{x} \in \mathbb{R}^k$,*

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(h_{n,1}(i) \leq x_i \sqrt{\ln n} + \ln n, i \in [k]\right) = \prod_{i=1}^k \Phi(x_i).$$

Proof. Recall the definition of $G_{n,x}$ in (8). We claim that for any $\bar{x} \in \mathbb{R}^k$,

$$(22) \quad \mathbf{P}\left(h_{n,1}(i) \leq x_i \sqrt{\ln n} + \ln n, i \in [k]\right) = (1 + o(1)) \prod_{i=1}^k \mathbf{E}[G_{n,x_i}(|\mathcal{S}_{n,1}(i)|)] + o(1).$$

To see this, let μ_n denote the law of $(\mathcal{S}_{n,1}(i), i \in [k])$ and ν_n denote the law of $(\mathcal{R}_n(i), i \in [k])$; recall that the latter are k independent copies of $\mathcal{S}_{n,1}(1)$. Let $\mathcal{B}_{n,k,1/2}$ be as defined in (17) and set

$$f_n(\bar{J}) = \mathbf{P}\left(h_{n,1}(i) \leq x_i \sqrt{\ln n} + \ln n, i \in [k] \mid \bar{\mathcal{S}}_{n,1} = \bar{J}\right),$$

$$g_n(\bar{J}) = \prod_{i=1}^k G_{n,x_i}(|J_i|).$$

From the first equation in Proposition 4.2, it follows that $f_n(\bar{J}) = g_n(\bar{J})$ for all $\bar{J} \in \mathcal{B}_{n,k,1/2}$. Therefore, the conditions on Lemma 6.1 for μ_n, ν_n and $\mathcal{B}_{n,k,1/2}$ are satisfied by Lemma 4.5 and Proposition 4.6, establishing (22).

Finally, it suffices to verify that, for all $i \in [k]$,

$$\lim_{n \rightarrow \infty} \mathbf{E}[G_{n,x_i}(|\mathcal{S}_{n,1}(i)|)] = \mathbf{E}\left[\Phi(\sqrt{2}x_i - N)\right] = \Phi(x_i);$$

where N is a standard Gaussian variable. The proof of this follows with the same argument as that for Lemma 3.3 with the main difference being that, instead of using $|\mathcal{S}_n(i)|$, we use $|\mathcal{S}_{n,1}(i)|$. By Lemma 4.1, the renormalization $\frac{|\mathcal{S}_{n,1}(i)| - 2 \ln n}{\sqrt{2 \ln n}}$ also converges to a standard Gaussian distribution. \square

We now proceed to treat the case with nontrivial conditioning.

Proposition 6.3. Fix an integer $k \geq 2$ and vectors $\bar{a} \in [0, 1]^k$, $\bar{b} \in \mathbb{Z}^k$ and $\bar{x} \in \mathbb{R}^k$; we have

$$\lim_{n \rightarrow \infty} \mathbf{P}(\mathbf{h}_{n,1}(i) \leq l_i, i \in [k] \mid \mathbf{d}_n(i) \geq m_i, i \in [k]) = \prod_{i=1}^k \Phi(x_i),$$

where $m_i = m_i(a_i, b_i, n) = \lfloor a_i \log n \rfloor + b_i$ and $l_i = l_i(a_i, x_i, n) = x_i \sqrt{\sigma_{a_i}^2 \ln n} + \mu_{a_i} \ln n$.

Proof. Recall the definition of $\tilde{G}_{m,l}$ in (12). In what follows, we assume, without lose of generality, that $m_i \geq 0$ (if $m_i < 0$ then $\mathbf{d}_n(i) \geq m_i$ a.s., so we set $m_i = 0$). Now, we first show that

$$(23) \quad \mathbf{P}(\mathbf{h}_{n,1}(i) \leq l_i, \mathbf{d}_n(i) \geq m_i, i \in [k]) = (1 + o(1)) 2^{-\sum_i m_i} \prod_{i=1}^k \mathbf{E} \left[\tilde{G}_{m_i, l_i}(|\mathcal{S}_{n,1}(i)|) \right] + o(1).$$

To see this, let μ_n denote the law of $(\mathcal{S}_{n,1}(i), i \in [k])$ and ν_n denote the law of $(\mathcal{R}_n(i), i \in [k])$; recall that the latter are k independent copies of $\mathcal{S}_{n,1}(1)$. Also, write

$$\begin{aligned} f_n(\bar{J}) &= \mathbf{P}(\mathbf{h}_{n,1}(i) \leq l_i, \mathbf{d}_n(i) \geq m_i, i \in [k] \mid \bar{\mathcal{S}}_{n,1} = \bar{J}) \\ g_n(\bar{J}) &= 2^{-\sum_i m_i} \prod_{i=1}^k \tilde{G}_{m_i, l_i}(|J_i|). \end{aligned}$$

Let $\alpha = \max\{a_i : i \in [k]\}$ and set $0 < \delta < 2 - \alpha$. Note that δ is chosen so that, for n large enough, $f_n(\bar{J}) = g_n(\bar{J})$ for all $\bar{J} \in \mathcal{B}_{n,k,\delta}$; this follows from Remark 4.3 and the second equation in Proposition 4.2. Lemma 4.5 and Proposition 4.6 yield the remaining conditions on μ_n, ν_n and $\mathcal{B}_{n,k,\delta}$, which applying Lemma 6.1 gives (23).

Next, let N be a variable with standard Gaussian distribution. For each $i \in [k]$,

$$(24) \quad \lim_{n \rightarrow \infty} \mathbf{E} \left[\tilde{G}_{m_i, l_i}(|\mathcal{S}_{n,1}(i)|) \right] = \mathbf{E} \left[\Phi \left(\frac{\sqrt{1 + \mu_{a_i}} x_i - N}{\sqrt{\mu_{a_i}}} \right) \right] = \Phi(x_i).$$

The last equality follows by Lemma 3.2 and the proof of the first equality follows similar to Lemma 3.3 when replacing the variables $|\mathcal{S}_n(i)|$ to $|\mathcal{S}_{n,1}(i)|$, which have the same limiting distribution.

Finally, it follows from (23) and (24) that

$$\lim_{n \rightarrow \infty} \mathbf{P}(\mathbf{h}_{n,1}(i) \leq l_i, \mathbf{d}_n(i) \geq m_i, i \in [k]) = 2^{-\sum_i m_i} \prod_{i=1}^k \Phi(x_i);$$

The result now follows by Proposition 5.4, since $\mathbf{P}(\mathbf{d}_n(i) \geq m_i, i \in [k])$ converges to $2^{-\sum_i m_i}$. \square

Proof of Theorem 2.5, case $k \geq 2$. Let $\bar{a} \in [0, 1]^k$ and $\bar{m} \in \mathbb{Z}^k$ be fixed and set $m_i = \lfloor a_i \log n \rfloor + b_i$. If $m_i \leq 0$ for all $i \in [k]$, the result follows from Proposition 6.2 and Lemma 5.2. Otherwise, the result follows from Proposition 6.3 and Lemma 5.5. \square

7. PROOF OF PROPOSITION 1.8

The next lemma appeared in [2]; we include its short proof for completeness.

Lemma 7.1. For any $k' \in \mathbb{N}$ and integers $(m_i, i \in [k'])$,

$$\mathbf{P}(\mathbf{d}_n(i) = m_i, i \in [k']) = \sum_{l=0}^{k'} \sum_{\substack{S \subseteq [k'] \\ |S|=l}} (-1)^l \mathbf{P}(\mathbf{d}(i) \geq m_i + \mathbf{1}_{[i \in S]}, i \in [k']).$$

Furthermore, for fixed $c \in (0, 2)$; if $m_i < c \ln n$ for $i \in [k]$ and $k' < k \in \mathbb{N}$, then

$$\begin{aligned} & \mathbf{P}(\mathbf{d}_n(i) = m_i, \mathbf{d}_n(j) \geq m_j, 1 \leq i \leq k' < j \leq k) \\ &= \sum_{l=0}^{k'} \sum_{\substack{S \subseteq [k'] \\ |S|=l}} (-1)^l \mathbf{P}(\mathbf{d}_n(i) \geq m_i + \mathbf{1}_{[i \in S]}, i \in [k]) \\ &= 2^{-k' - \sum_i m_i} (1 + o(1)). \end{aligned}$$

Proof. The first part is proven directly proved using the inclusion-exclusion principle. The second equation follows by intersecting the event $\{\mathbf{d}_n(j) \geq m_j, k' < j \leq k\}$ along all probabilities in the first equation; then applying Proposition 5.4 to each term:

$$\sum_{l=0}^{k'} \sum_{\substack{S \subseteq [k'] \\ |S|=l}} (-1)^l \mathbf{P}(\mathbf{d}_n(i) \geq m_i + \mathbf{1}_{[i \in S]}, i \in [k]) = (1 + o(1)) 2^{-\sum_i m_i} \sum_{l=0}^{k'} \sum_{\substack{S \subseteq [k'] \\ |S|=l}} (-1)^l 2^{-l}.$$

□

Corollary 7.2. Let $k' < k \in \mathbb{N}$ and fix $(a_i, A_i) \in \mathbb{Z} \times \mathcal{B}_I$ for $1 \leq i \leq k$. Write $m_i = \lfloor \log n \rfloor + a_i$ and

$$\begin{aligned} \mathcal{D}_{\bar{m}} &= \{\mathbf{d}_n(i) = m_i, 1 \leq i \leq k'\} \cup \{\mathbf{d}_n(i) \geq m_i, k' < i \leq k\}, \\ \mathcal{H}_{\bar{A}} &= \left\{ \frac{\mathbf{h}_n(i) - \mu \ln n}{\sqrt{\sigma^2 \ln n}} \in A_i, i \in [k] \right\}. \end{aligned}$$

Then

$$\mathbf{P}(\mathcal{D}_{\bar{m}}, \mathcal{H}_{\bar{A}}) = \left(2^{-k' - \sum_i d_i} \right) \prod_{i=1}^k \Phi(A_i) (1 + o(1)).$$

Proof. We start by intersecting the event $\mathcal{H}_{\bar{A}}$ along all probabilities in the second expression of Lemma 7.1; then we use the approximation by independent Gaussian variables given in Theorem 2.5. This gives

$$\begin{aligned} \mathbf{P}(\mathcal{D}_{\bar{m}}, \mathcal{H}_{\bar{A}}) &= \sum_{l=0}^{k'} \sum_{\substack{S \subseteq [k'] \\ |S|=l}} (-1)^l \mathbf{P}(\mathcal{H}_{\bar{A}}, \mathbf{d}_n(i) \geq m_i + \mathbf{1}_{[i \in S]}, i \in [k]) \\ &= \sum_{l=0}^{k'} \sum_{\substack{S \subseteq [k'] \\ |S|=l}} (-1)^l \mathbf{P}(\mathcal{H}_{\bar{A}} \mid \mathbf{d}_n(i) \geq m_i + \mathbf{1}_{[i \in S]}, i \in [k]) \mathbf{P}(\mathbf{d}_n(i) \geq m_i + \mathbf{1}_{[i \in S]}, i \in [k]) \\ &= (1 + o(1)) \prod_{i=1}^k \Phi(A_i) \sum_{l=0}^{k'} \sum_{\substack{S \subseteq [k'] \\ |S|=l}} (-1)^l \mathbf{P}(\mathbf{d}_n(i) \geq m_i + \mathbf{1}_{[i \in S]}, i \in [k]) \\ &= (1 + o(1)) \left(2^{-k' - \sum_i m_i} \right) \prod_{i=1}^k \Phi(A_i). \end{aligned}$$

□

Proof of Proposition 1.8. Fix $c \in (0, 2)$ and $M \in \mathbb{N}$. Let $j = j(n)$ and $j' = j'(n)$ be integer-valued functions with $0 \leq j'(n) + \log n < j(n) + \log n < c \ln n$; let $K' < K \in \mathbb{N}$ and $(a_k, k \in [K])$ be non-negative integers such that $\sum_{k \in [K]} a_k = M$ and set $M' = \sum_{k \in [K']} a_k$. Consider an arbitrary (K', K) -canonical sequence $((j_k, B_k), 1 \leq k \leq K)$ with $j' = j_1$ and $j = j_K$.

We define $m_i \in \mathbb{N}$ and $A_i \subset \mathbb{R}$ as follows. For each $k \in [K]$, if $\sum_{l=1}^{k-1} a_l < i \leq \sum_{l=1}^k a_l$ then set $m_i = \lfloor \log n \rfloor + j_k$ and let $A_i = B_{j_k}$. In this case, consider the sets

$$\begin{aligned} \mathcal{D}_{\bar{m}} &= \{d_n(i) = m_i, 1 \leq i \leq M'\} \cup \{d_n(i) \geq m_i, M' < i \leq M\}, \\ \mathcal{H}_{\bar{A}} &= \left\{ \frac{h_n(i) - \mu \ln n}{\sqrt{\sigma^2 \ln n}} \in A_i, i \in [M] \right\}. \end{aligned}$$

By Corollary 2.4 and the exchangeability of the vertex degrees of $T^{(n)}$,

$$\begin{aligned} \mathbf{E} &\left[\prod_{k=1}^{K'} \left(X_{j_k}^{(n)}(B_k) \right)_{a_k} \prod_{k=K'+1}^K \left(X_{\geq j}^{(n)}(B_k) \right)_{a_k} \right] \\ &= (n)_M \mathbf{P}(\mathcal{D}_{\bar{m}}, \mathcal{H}_{\bar{A}}) \\ &= (1 + o(1)) \left(2^{M \log n - M' - \sum_i m_i} \right) \prod_{i=1}^M \Phi(A_i), \end{aligned}$$

the last equality holding by Corollary 7.2 and since $(n)_M = n^M(1 + o(n^{-1}))$. Finally, note that

$$M \log n - M' - \sum_{i=1}^M m_i = \sum_{k=1}^{K'} (-j_k - 1 + \varepsilon_n) a_k + \sum_{k=K'+1}^K (-j' + \varepsilon_n) a_k,$$

and so

$$\left(2^{A \log n - A' - \sum_i m_i} \right) = \prod_{k=1}^{K'} (2^{-j_k - 1 + \varepsilon_n})^{a_k} \prod_{k=K'+1}^K (2^{-j' - \varepsilon_n})^{a_k}.$$

Similarly, $\prod_{i=1}^M \Phi(A_i) = \prod_{k=1}^K \Phi(B_k)^{a_k}$; which completes the proof. \square

8. PROOF OF THEOREM 1.1

Recall that \mathcal{M}_n is the set of vertices in t_n attaining the maximum degree. In light of Theorem 1.2, to prove Theorem 1.1 it suffices to prove the convergence of $|\mathcal{M}_n|$ over suitable subsequences.

Proposition 8.1. *Let $\varepsilon \in [0, 1]$. If n_l is an increasing sequence such that $\varepsilon_{n_l} \rightarrow \varepsilon$ as $l \rightarrow \infty$, then \mathbf{M}_{n_l} converges in distribution to \mathbf{M}_ε , where \mathbf{M}_ε is defined by*

$$\mathbf{P}(\mathbf{M}_\varepsilon = k) = \sum_{m \in \mathbb{Z}} e^{-2^{-m+\varepsilon}} \frac{2^{-(m+1-\varepsilon)k}}{k!}$$

for each integer $k \geq 1$.

Proof. The formula for $\mathbf{P}(\mathbf{M}_\varepsilon = k)$ may be seen using the following heuristic. Each of the terms in the sum represent the limit of the probability that; given that the maximum degree in T_n equals $\lfloor \log n \rfloor + m$, there exist precisely k vertices attaining such degree.

We first verify that $\sum_{k \geq 1} \mathbf{P}(\mathbf{M}_\varepsilon = k) = 1$; this follows from a telescopic analysis of the sum.

$$\begin{aligned} \sum_{k \geq 1} \mathbf{P}(\mathbf{M}_\varepsilon = k) &= \lim_{M \rightarrow \infty} \sum_{m=-M}^M \sum_{k \geq 1} e^{-2^{-m+\varepsilon}} \frac{2^{-(m+1-\varepsilon)k}}{k!} \\ &= \lim_{M \rightarrow \infty} \sum_{m=-M}^M e^{-2^{-m+\varepsilon}} \left(e^{2^{-(m+1-\varepsilon)}} - 1 \right) \\ &= \lim_{M \rightarrow \infty} \sum_{m=-M}^M \left(e^{-2^{-(m+1-\varepsilon)}} - e^{-2^{-m+\varepsilon}} \right) \\ &= \lim_{M \rightarrow \infty} \left(e^{-2^{-(M+1-\varepsilon)}} - e^{-2^{-M+\varepsilon}} \right) = 1 \end{aligned}$$

We now proceed to the proof of the theorem; we abuse notation by writing, e.g. $X_j = X_j(\mathbb{R})$. Consider $\varepsilon \in [0, 1]$ fixed and n_l an increasing sequence for which $\varepsilon_{n_l} \rightarrow \varepsilon$ as $l \rightarrow \infty$. We assume $\varepsilon = 0$ for simplicity of the formulas below. Fixing $k, M \geq 1$ we have

$$\mathbf{P}(\mathbf{M}_{n_l} = k) \leq \mathbf{P}(X_{\geq -M}^{(n_l)} = 0) + \sum_{j=-M}^{M-1} \mathbf{P}(X_j^{(n_l)} = k, X_{\geq j+1}^{(n_l)} = 0) + \mathbf{P}(X_{\geq M}^{(n_l)} > 0).$$

By Lemma 1.6 and Theorem 1.2 we have that for each $m \in \mathbb{N}$,

$$(X_{-m}^{(n_l)}, \dots, X_{m-1}^{(n_l)}, \dots, X_{\geq m}^{(n_l)}) \xrightarrow{\mathcal{L}} (X_{-m}, \dots, X_{m-1}, X_{\geq m})$$

and that the limit is a vector of independent vector of Poisson variables. In particular,

$$\mathbf{P}(X_j = k, X_{\geq j+1} = 0) = \mathbf{P}(X_j = k) \mathbf{P}(X_{\geq j+1} = 0) = \frac{e^{-2^{-j}} 2^{-(j+1)k}}{k!};$$

also $X_{\geq -M}^{(n_l)} = X_{\geq M}^{(n_l)} + \sum_{j=-M}^{M-1} X_j^{(n_l)}$, and $X_{\geq -M}^{(n_l)} \xrightarrow{\mathcal{L}} X_{\geq -M} \stackrel{\mathcal{L}}{=} \text{Poi}(2^M)$. Thus, in the limit

$$\begin{aligned} \limsup_{n_l \rightarrow \infty} \mathbf{P}(\mathbf{M}_{n_l} = k) &\leq \mathbf{P}(X_{\geq -M} = 0) + \sum_{j=-M}^{M-1} \mathbf{P}(X_j = k, X_{\geq j+1} = 0) + \mathbf{P}(X_{\geq M} > 0). \\ &= e^{-2^M} + \frac{1}{k!} \sum_{j=-M}^M \left(e^{-2^{-j}} 2^{-(j+1)k} \right) + \left(1 - e^{-2^{-M}} \right); \end{aligned}$$

This holds for arbitrary $M \in \mathbb{N}$, hence

$$\begin{aligned} \limsup_{n_l \rightarrow \infty} \mathbf{P}(\mathbf{M}_{n_l} = k) &\leq \liminf_M \left\{ e^{-2^{-(M)}} + \frac{1}{k!} \sum_{j=-M}^{M-1} \left(e^{-2^{-j}} 2^{-(j+1)k} \right) + \left(1 - e^{-2^{-M}} \right) \right\} \\ &= \frac{1}{k!} \sum_{j \in \mathbb{Z}} e^{-2^{-j}} 2^{-(j+1)k}. \end{aligned}$$

Similarly,

$$\liminf_{n_l \rightarrow \infty} \mathbf{P}(\mathbf{M}_{n_l} = k) \geq \limsup_M \left\{ \frac{1}{k!} \sum_{j=-M}^{M-1} \left(e^{-2^{-j}} 2^{-(j+1)k} \right) \right\} = \frac{1}{k!} \sum_{j \in \mathbb{Z}} e^{-2^{-j}} 2^{-(j+1)k}.$$

□

ACKNOWLEDGEMENTS

I would like to thank Louigi Addario-Berry and Henning Sulzbach for some very helpful discussions. This research was supported by FQRNT through PBEEE scholarship with number 169888.

REFERENCES

- [1] Louigi Addario-Berry. Partition functions of discrete coalescents: From cayley's formula to frieze's $\zeta(3)$ limit theorem. In *XI Symposium on Probability and Stochastic Processes: CIMAT, Mexico, November 18-22, 2013*, pages 1–45. Springer International Publishing, Cham, 2015.
- [2] Louigi Addario-Berry and Laura Eslava. High degree of random recursive trees. [arXiv:1507.05981](https://arxiv.org/abs/1507.05981), preprint (2016).
- [3] Nathanaël Berestycki. *Recent progress in coalescent theory*, volume 16 of *Ensaios Matemáticos [Mathematical Surveys]*. Sociedade Brasileira de Matemática, Rio de Janeiro, 2009.
- [4] François Bergeron, Philippe Flajolet, and Bruno Salvy. Varieties of increasing trees. In *CAAP '92 (Rennes, 1992)*, volume 581 of *Lecture Notes in Comput. Sci.*, pages 24–48. Springer, Berlin, 1992.
- [5] Shankar Bhamidi. Universal techniques to analyze preferential attachment tree and networks: Global and local analysis. <http://www.unc.edu/~bhamidi/preferent.pdf>, To be submitted to Probability Surveys.
- [6] Daryl J. Daley and David Vere-Jones. *An introduction to the theory of point processes. Vol. I. Probability and its Applications* (New York). Springer-Verlag, New York, second edition, 2003. Elementary theory and methods.
- [7] Daryl J. Daley and David Vere-Jones. *An introduction to the theory of point processes. Vol. II. Probability and its Applications* (New York). Springer, New York, second edition, 2008. General theory and structure.

- [8] Luc Devroye. Applications of the theory of records in the study of random trees. *Acta Inform.*, 26(1-2):123–130, 1988.
- [9] Luc Devroye and Jiang Lu. The strong convergence of maximal degrees in uniform random recursive trees and dags. *Random Structures Algorithms*, 7(1):1–14, 1995.
- [10] Michael Drmota. *Random trees*. SpringerWienNewYork, Vienna, 2009. An interplay between combinatorics and probability.
- [11] Richard Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, second edition, 1996.
- [12] William Goh and Eric Schmutz. Limit distribution for the maximum degree of a random recursive tree. *J. Comput. Appl. Math.*, 142(1):61–82, 2002. Probabilistic methods in combinatorics and combinatorial optimization.
- [13] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
- [14] Hosam M. Mahmoud. Limiting distributions for path lengths in recursive trees. *Probab. Engrg. Inform. Sci.*, 5(1):53–59, 1991.
- [15] Tamás F. Móri. The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.*, 14(3):339–348, 2005.
- [16] Hwa Sung Na and Anatol Rapoport. Distribution of nodes of a tree by degree. *Math. Biosci.*, 6:313–329, 1970.
- [17] Boris Pittel. Note on the heights of random recursive trees and random m -ary search trees. *Random Structures Algorithms*, 5(2):337–347, 1994.

DEPARTMENT OF MATHEMATICS AND STATISTICS, MCGILL UNIVERSITY,
 805 SHERBROOKE STREET WEST, MONTREAL, QUEBEC H1X 2B5, CANADA
E-mail address: `laura.eslavafernandez@mail.mcgill.ca`
URL: `http://www.math.mcgill.ca/eslava/`